DOCUMENT RESUME

ED 416 227 TM 028 102

AUTHOR Ramsey, Sarah H.; Kelly, Robert F.
TITLE Using Social Science Research in Family Law Analysis and Formation: Problems and Prospects.
PUB DATE 1998-00-00
NOTE 91p.
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Cooperation; Decision Making; *Educational Policy; Educational Research; Family Programs; *Lawyers; Legislation; *Policy Formation; Public Policy; *Research Utilization; *Researchers; *Social Science Research; Social Sciences
IDENTIFIERS *Family Law; Reform Efforts

ABSTRACT

Social science research can make a valuable contribution to family law analysis and formation. It can help define problems, identify possible solutions, and challenge underlying normative assumptions. Recent studies related to family law reform have analyzed the use of wage-withholding and other changes to increase child support amounts and collection, the impact of divorce on children, and parental decision making about custody on divorce. To encourage the responsible use of social science research for policymaking, including educational policymaking, this article explores the limits and potential of law and social science research interaction. Part II of this article explores these limitations by describing four major methodological problems with which family research, as a type of social science, must deal: (1) measurement; (2) sampling and generalizing; (3) problems of causal inference; and (4) researcher bias. Included in this discussion is a detailed assessment of the problem of bias in the research process and recent feminist contributions to the analysis of bias. It also discusses recent advances in the quality of family policy research. Part III deals with the use of research for policy analysis and formation and explains the need for multiple studies and the normative nature of policymaking. It also analyzes possible roles for comprehensive reviews of research literature in the policy process. The conclusion of the article emphasizes that developing criteria for the use of family social science research in law formation is a formidable undertaking. A proposal is presented for the collaboration of social science researchers and lawyers to improve this relationship. The proposed panel of lawyers and researchers would not be a funding agency, but would be a structured forum for discussion of the family policy process. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

#93-23

ED 416 227

Using Social Science Research in Family Law Analysis and Formation: Problems and Prospects*

by

Sarah H. Ramsey** and Robert F. Kelly***

*The authors are grateful for the many helpful suggestions made by Peter Bell, Mary Jo Coiro, Michael Wald and other colleagues on earlier drafts of this article and the student research assistance of Paul Schmidtberger (Stanford Law School) and David Efron (Syracuse). In addition we appreciate the support provided by Stanford Law School, Syracuse University College of Law, LeMoyne College, Child Trends, Inc. (Washington, D.C.) and the William and Flora Hewlett Foundation.

** Professor of Law and Associate Dean, Syracuse University College of Law; B.A., Duke University, 1965; J.D. University of North Carolina at Chapel Hill, 1973; LL.M., University of Michigan, 1982.

***Professor of Sociology and Chair, LeMoyne College; B.A. (1972), St. Joseph's University; M.A. (1975) and Ph.D. (1978), Rutgers University.

TM028102

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Carol Emig

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

Table of Contents

I. Introduction

II. The Limitations of Social Science Research: Methodological Problems

A. Measurement

B. Sampling and Generalizing

C. Problems of Casual Inference

D. Researcher Bias

1. Claims of Objectivity

2. The Existence of Bias: Feminist Critiques

3. What Can Be Done about Bias?

III. Using Research in Family Law Analysis and Formation

A. The Need for Multiple Studies and the Normative Basis of Policymaking

B. The Research Review

IV. Conclusions and a Proposal

I. Introduction

Social science research can make a valuable contribution to family law analysis and formation. It can help define problems, identify possible solutions and challenge underlying normative assumptions. Recent studies related to family law reform, for example, have analyzed the use of wage-withholding and other changes to increase child support amounts and collection,[1] the impact of divorce on children,[2] the impact of foster care on children,[3] and parental decision making about custody upon divorce.[4] Although the ultimate choice of a policy is a normative decision, not something any of these studies could determine, research can inform and improve the quality of the policy debate and public discourse that leads up to law reform.

Unfortunately, recent criticism of the use of social science

---

[1] Irwin Garfinkel et al., The Wisconsin Child Support Assurance System, 24 J. of Hum. Resources 1 (1990) [hereinafter Garfinkel 1990]; Irwin Garfinkel & Marieka M. Klawitter, The Effect of Routine Income Withholding of Child Support Collections, 9 J. of Pol'y Analysis and Mgmt. 155 (1990); Irwin Garfinkel, Utilization and Effects of Immediate Income Withholding and the Percentage-of-Income Standard: An Interim Report on the Child Support Assurance Demonstration, 4 Inst. for Res. on Poverty, Special Report No. 42 (1986) [hereinafter Garfinkel 1986].

[2] Andrew J. Cherlin et al., Longitudinal Studies of the Effects of Divorce on Children in Great Britain and the United State, 252 Science 1386 (1991).

[3] Michael S. Wald Et Al., Protecting Abused and Neglected Children (1988).

[4] Eleanor Maccoby & Robert H. Mnookin, Dividing the Child: Social and Legal Dilemmas of Custody (1992). Robert H. Mnookin et al., Private Ordering Revisited: What Custodial Arrangements Are Parents Negotiating? in Divorce Reform at the Crossroads (Sugarman & Kay, eds. 1990).

BEST COPY AVAILABLE

research in family law may discourage legal scholars, law students and social scientists from undertaking the difficult, but needed, interdisciplinary task of relating social science and law. The critics have faulted both the researchers and the users of the research.[5] The researcher is criticized for doing biased, inaccurate research, and for presenting it in a misleading manner. The critics suggest also that the research is misunderstood and used selectively and inappropriately by law reformers. One critic even suggests that the influence of social scientists "has led to apparent reduction in the intellectual challenge and content of the law."[6]

To encourage the responsible use of social science research for policymaking, this article explores both the limits and potential of law and social science research interaction.[7] This article provides a primer for the legal scholar or student who

---

[5] Ruth Deech, Divorce Law and Empirical Studies, 106 Law Q. Rev. 229 (1990); David L. Faigman, To Have and Have Not: Assessing the Value of Social Science to the Law as Science and Policy, 38 Emory L.J. 1005 (1989); Martha L. Fineman and Anne Opie, The Uses of Social Service Data in Legal Policymaking: Custody Determinations at Divorce, 1987 Wis. L. Rev. 107 (1987); Martha L. Fineman, Custody Determination at Divorce: The Limits of Social Science Research and the Fallacy of the Liberal Ideology of Equality 3 Canadian J. of Women and the Law 88 (1989) [hereinafter Fineman, 1989]; Martha A. Fineman, The Illusion of Equality: The Rhetoric and Reality of Divorce Reform (1991) [hereinafter Fineman, 1991].

[6] Deech, supra note 5, at 245.

[7] Although courts also make policy, this Article does not address the problems in using social science as evidence. For interesting proposals on this subject see Laurens Walker & John Monahan, Social Frameworks: A New Use of Social Science in Law, 73 Va. L. Rev. 559 (1987).

BEST COPY AVAILABLE

wants to use social science research in legal analysis or law formation. We emphasize that users of research involved with analyzing existing law or recommending a reform need to be concerned about the quality of the research they are using, the quality of their own analysis and summary of the research, and the nature and degree of reliance that they then place on their research review.

To emphasize the need for caution, Part II of this article explains the limitations of social science research by describing the four major methodological problems with which family research, as a type of social science, must deal. Included in this discussion is a detailed assessment of the problem bias in the research process and recent feminist contributions to the analysis of bias. It also discusses recent advances in the quality of family policy research. Part III deals with the use of research for policy analysis and formation and explains the need for multiple studies and the normative nature of policymaking. It also analyzes possible roles for comprehensive reviews of research literatures in the policy process.

The conclusion of this article emphasizes that developing criteria for the use of family social science research in law formation is a formidable undertaking. Although achieving a consensus on what and how research can be used is unlikely in the near future, concrete steps can be taken to enhance the quality of the relationship between family social science and family law analysis and formation. Toward this end, we describe a proposal

illustrating how organizations of lawyers and social scientists might collaborate to improve this relationship. Though social science research cannot resolve family law reform debates, it can significantly improve the quality of the debates.[8] However, much work remains if this potential is to be realized.

II. The Limitations of Social Science Research: Methodological Problems

In this Part we describe and analyze in a general way the four fundamental methodological problems that the social scientist doing family research confronts: measurement, sampling and generalizability, causal inference, and researcher bias. Considering these four basic problems will provide sufficient evidence for the need for caution in using social science, although there are numerous other technical methodological issues that we might also have addressed in a more extended analysis.[9]

---

[8] For an interesting discussion of the Supreme Court's use of empirical research and the value of research as a constraint on the court, see David L. Faigman, Normative Constitutional Fact-Finding: Exploring the Empirical Component of Constitutional Interpretation, 139 U. Pa. L. Rev. 541 (1991). For a general discussion of the practical contributions made by the social sciences to the quality of life in contemporary society see Kenneth Prewitt, Usefulness of the Social Sciences, 211 Science 659-(1981).

[9] Some useful, general texts on social science research methods are John B. Williamson, et al., The Research Craft: An Introduction to Social Research Methods, [hereinafter Research Craft]; Earl R. Babbie, The Practice of Social Research (1989). Two useful sources on policy research are Peter H. Rossi & Howard E. Freeman, Evaluation: A Systematic Approach (1982), and James S. Coleman, Problems in Conceptualization and Measurement in Studying Policy Impacts, in Public Policy Evaluation 19 (Kenneth M. Dolbeare, ed. 1975).

BEST COPY AVAILABLE

A. Measurement

In designing a study to test hypotheses related to families,[10] deciding what to measure and how is a complex and challenging task. Consider for instance, just the problems involved in defining the major policy or outcome variables for a study of the impact of divorce on children.[11] Should the effects of "divorce" be measured from the date of the actual court decree of divorce, or an earlier date such as the date of separation, or the date of the onset of significant marital conflict and disruption?[12] If we wanted to determine whether children from families that had not experienced marital

---

[10] See Karl Popper, Objective Knowledge: An Evolutionary Approach 13-17 (1972). For a succinct discussion of the concept of falsifiability and theory testing, see Faigman, supra note 5 at 1016-21. An hypothesis is a theoretically grounded expectation that some independent variable, the suspected cause, is systematically related to some dependent variable, the suspected effect. The scientific method requires that we test, i.e. seek to falsify, hypotheses against empirical reality.

[11] A policy variable is a type of independent variable amenable to policy control that a policymaker or proponent of a policy believes will have a beneficial impact on a policy outcome variable, a type of dependent variable. Coleman, supra note 9 at 24-25. The expectation that a policy variable will influence a policy outcome variable in the anticipated manner may be thought of as a policy hypothesis. For example, some proponents of joint-legal custody (the policy variable) suggested, i.e., hypothesized, that it would reduce post-divorce litigiousness and enhance co-parenting subsequent to divorce (two policy outcome variables).

[12] Researchers have used different definitions. Cherlin et al., supra note 2 (date of separation); Judith S. Wallerstein & Joan R. Kelly, Surviving the Breakup: How Children and Parents Cope with Divorce (1980) [hereinafter Surviving the Breakup] (Parents in the sample had separated and filed for divorce; although the authors discuss measuring the effects on the child if the parents divorce, they appear to have measured effect without regard to the actual date of divorce).

BEST COPY AVAILABLE

disruption fared better than children who had experienced divorce, how should we define "better"? This is not an issue that child development research can resolve, because "better" is a normative concept.[13] Should we primarily be concerned with emotional well-being or educational attainment, or both, for example? Are we interested in short term or long term assessment? Are we more concerned, for example, about how the child fares the first year after divorce, five years after divorce, or do we want a series of measures over time that will somehow be cumulative?

In doing policy research part of our answer to these measurement questions would be controlled by practical concerns related to the nature of our audience.[14] An expansive list of multiple outcome variables would probably be needed to respond to the interests of various constituent groups. Some might be more interested in health related concerns, for example, and some in education. Since we would not want our results to be rejected on the basis that we were measuring the "wrong" aspects of the child's development, we should anticipate challenges to our decisions about what to measure, about what constitutes "better," and we should be as comprehensive as possible in measuring child outcomes.

---

[13] Robert H. Mnookin, In the Interests of Children 18 (1985) ("Deciding what is best for a child often poses a question no less ultimate than the purposes and values of life itself.")

[14] See, e.g., Research Craft, supra note 9, at 27-28.

BEST COPY AVAILABLE

Strategies for measuring the impacts of family law change or innovations that are inclusive and comprehensive with respect to the concerns of multiple interest groups and constituencies have the added advantage that they are also consistent with the principles of measurement theory. Researchers have long worked with the dual notions that no single indicator of a concept, for example the well-being of children, will be adequate to empirically capture the fullness of the reality of such a concept and, at the same time, that all individual indicators of concepts will contain some amount of measurement error.[15] The implication of this understanding is that measurement validity, that is the ability of a measurement strategy to measure what it intends to measure, optimally requires the use of multiple and independently derived indicators of the concept to be measured.[16] This strategy does not assure perfect measurement--there is no such thing as perfect measurement--but rather it maximizes the likelihood that measurement error will be minimized. Hence, the

---

[15] On the need to use multiple measures collected from a variety of independent sources see Research Craft, supra note 9, at 81-83; Edward J. Carmines and Richard A. Zeller, Reliability and Validity Assessment 29-30, 32 (1979) and John L. Sullivan and Stanley Feldman, Multiple Indicators: An Introduction 16 (1979); Rossi and Freeman make the same point in their text on public policy and evaluation research. See Rossi & Freeman, supra note 9, at 151-60.

[16] For a standard discussion of measurement validity see Carmines and Zeller supra note 15 at 11-13. On the need for multiple questions measuring the same concept in survey research see Jean M. Converse and Stanley Presser, Survey Questions: Handcrafting the Standard Questionnaire 45,47 (1986).

BEST COPY AVAILABLE

measurement strategies used in research relevant to family law analysis and formation should be judged both with respect to the policy imperative that the interests of multiple groups be represented, and with respect to the measurement theory imperative that multiple independent indicators be used to minimize error and enhance measurement validity.

In his study of the impact of foster care on abused and neglected children, Michael Wald and his colleagues demonstrated an appreciation and understanding of these measurement issues in their use of multiple outcome indicators.[17] Because the goal of the research was to provide information to policymakers about intervention strategies, the researchers wanted to have measures of the child's well-being that policymakers would consider relevant. They assumed that "legislators and judges are concerned with physical health and evidence of malfunctioning, such as poor school performance, severe emotional disorders, or antisocial behavior...."[18] Consequently the researchers measured these factors, as well as others they thought were important, such as the child's own perspective on the intervention. The study was longitudinal with assessments being made periodically over a two year period. Multiple measures and independent sources were used for these assessments. To measure the children's health, for example, a physical examination was given at the beginning and end of the study, height and weight

---

[17] Wald, supra note 3.

[18] Id. at 28.

assessments were made annually, and caretakers were asked to rate the child's health and to provide a medical history.

To appreciate the importance of multiple independent indicators of a concept such as the health of a child, it is instructive to compare Wald's approach with a measurement strategy for assessing well-being that relies exclusively on a single indicator such as the number of visits per year to the doctor. This indicator may measure health, but it may also measure how affluent the child's parents are and their access to health care. Hence, if this indicator was relied upon alone as a measure of well-being, measurement error would likely be introduced into the researcher's analysis.

Although measurement validity is enhanced by using multiple independent measures of a concept, practical considerations may limit their use. There are limits to the time commitment that can be expected from study participants before the quality of the information they provide deteriorates. Further, costs and other constraints typically prevent researchers from fully implementing their ideal data collection strategies. For example, in their analysis of the impact of divorce on children, Cherlin and his colleagues constructed scales of children's behavior problems based only on interviews with the parent most familiar with the child's daily activities, usually the mother's, at two points in time for the United States portion of their cross-national

analysis.[19] This is not an optimal measurement strategy because we know that adults in different roles see children in different ways. Parents, because they have known their children for the entire duration of the child's life, are more likely to detect changes in attitudes and behavior, and to identify problems such as depression and withdrawal. Teachers, on the other hand, because of the demands of classroom management, will be more sensitive to problems such as acting out and attention disorders.[20] In a study of divorce, these measurement issues are complicated even further by the fact that parents may have an investment in perceiving that their children are adjusting well to the divorce and its consequences. Clearly, confidence in the Cherlin et al findings would have been enhanced if the researchers had also been able to use measures of behavioral problems derived from multiple sources such as teachers and mental health counselors as well as parents.

The Cherlin et al study reflects the practical difficulties and the necessary compromises related to measurement that characterize even the most rigorously designed and implemented family study. The U.S. data set used by Cherlin, the National Survey of Children, a nationally representative ten-year longitudinal survey of American children, actually included

---

[19] Cherlin et al., supra note 2, at 1388. Cherlin's study used data from the first two waves of the three waves of the National Survey of Children.

[20] Thomas M. Achenbach et al., Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity, 101 Psychological Bulletin 213 (1987).

interviews with the teachers of the children in the study, but unfortunately the information gathered from teachers mainly concerned academic functioning and achievement, and developmental maturity, rather than behavioral problems. Hence, information on the child, while collected from multiple independent sources, concerned different although equally important topics. Thus, the authors could not combine the information provided from these different perspectives for purposes of measuring behavioral problems. In addition, because the analysis was comparing the impact of divorce on American and British children, the researchers were constrained to focus their analysis on child behavioral measures that were common to the data from both data sets.[21] In spite of these difficulties with regard to multiple independent sources of measurement, the longitudinal and nationally representative nature of the National Survey of Children and its rigorous analysis, make the Cherlin et. al. study a significant advance in family research.

It is also important to note that measurement validity is enhanced when researchers employ measures that have been standardized and successfully employed in other studies. The social scientific enterprise of measurement is an ongoing and historical process in which current researchers are expected to capitalize upon both the achievements and the mistakes of past

---

[21] Based on personal communications and background working papers provided by Donna Ruane Morrison of Child Trends, Inc.(Washington, DC), (one of the coauthors of Cherlin's study of the effects of divorce on children in Great Britain and the United States, supra note 2).

BEST COPY AVAILABLE

researchers. Hence, in evaluating measurement strategies used in policy relevant family research, it is important to ask about the history of the measures being used and the degree to which particular measures have been critically evaluated and accepted by the scientific community. This understanding of measurement does not imply that only established measures should be used, but rather that to be accepted new measures need to survive a period of scrutiny in which they are compared with established measures as to their efficacy. In this regard it is notable that although the Cherlin analysis only employed interviews with a parent to assess the impact of divorce on children's behavior, the series of child behavioral outcome questions that they used were short versions of widely accepted and standardized measurement instruments (Behavioral Problem Index) commonly used in many other studies.[22]

In addition to measurement validity, researchers and consumers of research also need to be concerned with measurement reliability. When researchers speak of the reliability of a measure or scales, for example a question or series of questions concerning the well-being of children, they are referring to the degree to which the measure or scale can be expected to yield stable results across repeated research trials or surveys when administered under similar circumstances.[23] We do not want to

---

[22] Id.

[23] For general treatment of reliability see Research Craft, supra note 9, at 70-71; See also Babbie, supra note 9, at 121-24; and Rossi & Freeman, supra note 9, at 173-74; and Carmines &

BEST COPY AVAILABLE

use a yard stick that gives different measures of the same distance depending on the temperature or who is using it.

Reliability can frequently be a problem in family research. For example, Wald was concerned that in responding to questions about their children's behavior, the abusive or neglectful parents in his study might view a child's crying or whining as more problematic than would a non-abusive parent and therefore would rate their children differently. He also was concerned that foster mothers, who were paid to deal with the children, might be less likely to categorize certain behaviors as problems. The question here is: when similar questions are asked about the actions of children who are behaving similarly, is the researcher getting similar responses or readings? Further, researchers using large national surveys to study the impact of divorce and family structure on child well-being have been concerned about the reliability of child behavioral problem scales based on reports from biological mothers and fathers, stepmothers and fathers, and custodial versus non-custodial mothers and fathers. Fortunately, statistical techniques have been developed that allow researchers to quantitatively assess the level of reliability of measures and scales. As in the case of measurement validity, those seeking to use family social science in family law analysis and formation need to be aware of and critically assess the reliability of the measures used in the

---

Zeller, supra note 14 at 11-13.

studies they employ.[24]

Another measurement decision the researcher must make is whether to collect quantitative data, qualitative data, or both. Social scientists typically use the term qualitative research to refer to methodologies such as collecting information from long and intense, but relatively unstructured, interviews, or conducting long-term observational studies such as those used by anthropologists to produce deeply descriptive accounts of collective symbol systems and behaviors.[25] Quantitative research typically refers to the collection of information through highly structured and standardized experimental and survey designs, and the analysis of this information with powerful statistical techniques such as multivariate modeling.[26]

An advantage of quantitative data collection techniques such as surveys is that, when carried out properly, the classification

---

[24] For discussions of techniques used to assess reliability see Hubert M. Blalock, Jr., The Measurement Problem: A Gap Between the Languages of Theory and Research in Methodology in Social Research (Hubert M. Blalock and Ann B. Blalock eds. 1968) [hereinafter Methodology in Social Research]; Paul M. Siegle and Robert W. Hodge, A Casual Approach to the Study of Measurement Error in Methodology in Social Research; and Carmines & Zeller, supra note 15 at 37-51.

[25] Two basic texts on the rationale for and the use of qualitative methods in social sciences are Robert Bogdan and Steven Taylor, Introduction to Qualitative Research Methods: A Phenomenological Approach to the Social Sciences (1975); and John Lofland and Lynn H. Lofland, Analyzing Social Settings: A Guide to Qualitative Observations and Analysis (2d ed. 1984). For a highly regarded example of qualitative research methodology used in a study of families and public policies see Carol B. Stack, All Our Kin: Strategies for Survival in a Black Community (1974).

[26] For a general discussion of their approach see Paul E. Spector, Research Designs (1981).

BEST COPY AVAILABLE

and interpretation of the data are less varied because they use standardized questions with categorized answers that are administered by data collectors trained to treat each interviewee in a consistent manner. As a result, it is easier for other researchers to replicate and critique the study. In contrast, the collection and analysis of qualitative data are more subjective and interpretive because of the researcher's high level of involvement with the subjects of the research, making a critical review more difficult.

On the positive side, qualitative data can provide a more comprehensive, nuanced and complex picture of social events, processes and relationships. When a researcher is asking an intimate question, such as why couples separate, an intimate methodology of the type employed by anthropologists that can follow people closely over time is appropriate.[27] Indeed, the ability of qualitative research to observe, document and interpret social processes is its major advantage over quantitative methods that typically measure outcomes and statuses more than processes. Using both qualitative and quantitative data in family research often represents the ideal measurement strategy because the data sets can be analyzed separately and in combination, allowing the strengths of each type of data to

---

[27] Robert F. Kelly, Family Policy Analysis: The Need to Integrate Qualitative and Quantitative Research Methods, 13 Soc. Methods and Res. 363 (1985). For more detailed discussion with extensive research examples, see Qualitative Methods in Family Research (Gilgun, Daly & Handel, eds. 1992).

BEST COPY AVAILABLE

improve the analysis.[28]

In summary, family research is improved by the use of multiple measures, independent data sources, standardized measures, assessments of reliability, and a variety of data collection techniques. As studies become more complex, however, data collection, analysis and interpretation become more difficult and expensive. The extensive measures that Wald used give us a fuller picture of the well-being of the children studied than does the Cherlin study. Wald's study, however, collected data on only 32 abused and neglected children but cost over 1.5 million dollars and took over 8 years to complete.[29] Collecting extensive data from multiple sources on various

---

[28] Id. Kelly used the example of the Seattle-Denver Income Maintenance Experiment (SIME/DIME). The SIME/DIME researchers had expected to find that in comparison to AFDC, the negative income tax would be associated with higher marital stability because two parent families were as eligible as one parent families. Instead, they found the opposite: the rate of separation and divorce was higher for the negative income tax recipients than for the AFDC recipients. Kelly pointed out that the SIME/DIME methodology was not suited to answer the question of why the negative income tax recipients were more likely to separate or divorce, because it focused on the negative income tax as a cause and looked at its impact, without looking at why it had certain effects. He suggested that to answer the "why" question about marital instability, an intimate question, an intimate methodology would have been appropriate. Kelly also points out, however, that families that were intensively interviewed or observed might be omitted from the quantitative analysis because their responses might be affected by the intense interview process.

For additional research examples of the benefits of the combination of both types of data see Mark R. Rank, The Blending of Qualitative and Quantitative Methods in Understanding Childbearing Among Welfare Recipients, in Qualitative Methods in Family Research, supra note 27 at 281; and Margonete Sandelowski et. al., Using Qualitative and Quantitative Methods: The Transition to Parenthood of Infertile Couples, in id. at 301.

[29] Wald et al., supra note 3, at 216, n. 23.

BEST COPY AVAILABLE

measures for large samples is work that clearly needs to be undertaken, but the cost of doing so is substantial.[30]

B. Sampling

A key issue to be considered in evaluating social science research for family law analysis and policy formation is sampling. Because of cost considerations, social scientists typically sample members of populations rather than studying entire populations.[31] A major challenge in family research is developing representative samples that can used to generalize to the larger populations of interest. Policy research directed toward state law formation, for example, should have a sample that reflects the characteristics of the relevant population in that state. If this population has people of different socio-economic and ethnic backgrounds, the sample should also, preferably in similar proportions.

---

[30] Kirstin Moore, Executive Director of Child Trends, Inc., the original contractor of the National Survey of Children, estimates that a new national survey of 10,000 children with assessments at three points in time would cost between ten and twenty million dollars depending on the degree to which telephone interviews were used, the amount of direct observation of children in the field and the extent to which multiple sources of information are gathered for each child (personal communication). It is notable that Canada has recently undertaken a national longitudinal survey of children of this type (correspondence with the Social Program Information Division, Health and Welfare Canada).

[31] Note that if cost is not an issue, that is, if information on an entire population is available at a low cost and if analysis will not be slowed because one has chosen to study an entire population, then studying the entire population is clearly preferable to studying a sample of the population.

The optimal method for achieving a representative sample is to use an unbiased random function to select members of the sample from a universal listing of all members of the population that the sample is to represent.[32] A sample selected in this manner is called a simple random sample. There are many technical variations of the simple random sample (generally referred to as probability samples), but the key notion for the consumer of family social scientific research to keep in mind is that non-probability samples may be biased, and are therefore likely to present a distorted picture of the population. In contrast, well designed and implemented probability samples are said to possess high levels of external validity or generalizability.

Producing a representative probability sample can be expensive and difficult. Respondents chosen randomly will typically be diffused across the entire area inhabited by the population, rather than concentrated in areas easy for the researcher to access. Once subjects are selected for the sample, it is crucial that the researcher maximize the participation rate of these individuals or households or families, if the benefits of a random sample are to be realized. Securing a high participation rate typically involves time-consuming follow-up efforts and, in some instances, costly inducements for

---

[32] On sampling generally see Babbie, supra note 9, at 161-207; and Research Craft, supra note 9, at 103-22. For a thorough but technical discussion see Leslie Kush, Survey Sampling (1965) and Graham Kalton, Introduction to Survey Sampling (1983). Computer programs and random number charts available in most statistics texts provide random functions that can be used in sampling procedures.

BEST COPY AVAILABLE

prospective subjects. If the study is longitudinal, maintaining contact with subjects in the original random sample can be an expensive and daunting task. For these reasons, and unfortunately because rigor in sampling is not uniform in the social sciences, many researchers rely on small, non-random samples in their work. For example, in their much cited work on the impact of divorce on children, Judith Wallerstein and Joan Kelly used a sample of families with children experiencing divorce who presented themselves to a family center for short-term divorce related counseling and agreed to participate in the study.[33] This type of sample is often referred to as a convenience sample because it capitalizes on social situations where groups of interest to researchers are concentrated, in this case a clinical setting. Another frequently used type of convenience sample is the snowball sample.[34] In snowball sampling volunteers are solicited from likely sources and then asked to identify others like themselves for inclusion in the sample.[35]

---

[33] Surviving the Breakup, supra note 12, app. at 319-34 (Method and Sample).

[34] Babbie, supra note 9, at 268-69.

[35] This technique was used by Santrock and Warshak to identify father-custody families in their study of the impact of the custodial parent's gender in the child's functioning. John W. Santrock & Richard A. Warshak, Father Custody and Social Development in Boys and Girls, J. Soc. Issues (Fall 1979). One of the strongest justifications for employing convenience sampling occurs when the population one wishes to study is very small as a proportion of the total population and very difficult to locate. Fathers with sole physical custody of minor children clearly qualify as such a group.

BEST COPY AVAILABLE

The disadvantages of these convenience samples are that they are likely to be biased, that is not representative of the population to which the researcher wishes to generalize, and there is no way to assess the nature or the degree of the probable sampling bias with any degree of certainty. Hence, findings based on analyses of these samples cannot and should not be generalized to other groups or the larger population with any degree of confidence.[36] Divorcing families that have gone to a counseling center for help, for example, may have different and/or more severe problems than would a randomly selected group of families experiencing divorce. On the other hand, people who seek out counseling may be healthier than the general population of divorcing families as indicated by the resourcefulness they demonstrate in seeking help. The problem is that we cannot know how these possible biases may influence results if we are working with a nonprobability convenience sample.

With this understanding of sampling as a background, it is important to recognize that some family researchers have been much less cautious than they should be in generalizing from convenience samples. Wallerstein and Kelly, for example, studied a very limited sample of 131 children from 60 predominantly white, upper middle and middle class families that sought services at a Marin County, California counseling center, but suggested that their finding that divorce has a permanent, long-

---

[36] Research Craft, supra note 9, at 104; and Kaltan, supra note 32 at 90.

term negative effect on children was likely to have wide applicability.[37] Cherlin's study, that was based on two nationally representative surveys, one from the United States and one from Great Britain, does not support Wallerstein and Kelly's dire conclusions but rather suggests that many of the problems children of divorce experience are related to conditions that existed well before the event of their parent's separation.[38] On sampling grounds then, we should have greater confidence in our ability to generalize from Cherlin's findings than those of Wallerstein and Kelly.

Identifying large random samples of specific family configurations such as divorcing or divorced couples can be particularly difficult and expensive. Some researchers studying divorce have used court records to locate their research subjects. In their study of custodial arrangements, for example, Maccoby and Mnookin went through 7000 court records in two California counties to identify 1966 families that met their criteria for inclusion in the study. Slightly less than sixty percent of these (1124) were located and agreed to participate.[39] Other researchers have been less successful using this approach, however, and tracked down so few families

---

[37] Surviving the Breakup, supra note 12, at 312-17.

[38] See Cherlin et al., supra note 2, at 1388.

[39] Maccoby & Mnookin, supra note 4; Mnookin et al., supra note 4.

that the representativeness of their samples is questionable.[40] A low response rate is always a source of concern in using family social science, even if the study employs a random sample, because if potential members of the sample refuse to participate and they share some underlying characteristic that is relevant to the phenomenon being studied, the results of the study can be biased. Fortunately, this problem to some degree can be remedied by a variety of techniques that have been developed to adjust for nonresponse in random sample surveys. Consumers of family social science should be concerned about sample response rates and aware of the importance of researchers using adjustment techniques for nonresponse.[41]

An additional problem with identifying families for study at a point in time after an event such as divorce is that information about how the child and family were functioning prior to divorce would be retrospective and therefore likely to be less accurate.[42] If we wanted to study the impact of divorce on

---

[40] Another study relying on court records, located only 37% of the potentially eligible respondents and interviewed less than 25%. Frank Furstenberg & Graham Spanier, Recycling the Family: Remarriage After Divorce 16-17 (1987 updated ed.) (describing Spanier's 1976 study in Centre County, Pennsylvania).

[41] For a general introduction to these problems see Richard A. Berk, An Introduction to Sample Selection Bias in Sociological Data, 48 Am. Soc. Rev. 386-98 (1983) and Kalton supra note 32, at 63-69. Weighting procedures and the use of instrumental variables, derived from nonresponse selection models, in multivariate analyses are common approaches to nonresponse in sampling.

[42] See, e.g., Paul R. Amato, The "Child of Divorce" as a Person Prototype: Bias in the Recall of Information About Children in Divorced Families, 53 J. of Marriage and the Fam. 59 (1991).

BEST COPY AVAILABLE

children's health, for instance, we would need to know both whether a particular child had certain chronic health problems prior to divorce and the level of conflict that characterized the marriage in the years leading up to the divorce so that we could control for these preexisting factors and avoid wrongly attributing their effects to divorce. If the study did not begin until after divorce, we would be asking parents and others to recollect past events and behaviors, recollections likely to be influenced by the contemporary fact of divorce.

An alternative to using samples of convenience and court records to identify divorced couples that avoids the problem of retrospective data collection is using a longitudinal, general population survey that collected data on the same families at multiple points in time.[43] Over time we could reasonably expect that a sizable proportion of families in such a sample would separate and divorce so that the researcher would have information about the families that was collected both before and after the marital disruption.

In addition to allowing for before-after comparisons of families that experience divorce, longitudinal designs have the advantage of creating a comparison or control group of families who did not experience divorce during the period studied. As we shall see in the next section, the ability of researchers to make causal inferences, in this case concerning the impact of events

---

[43] On longitudinal designs see Babbie, supra note 9, at 89-91; and Research Craft, supra note 9, at 129-30.

such as divorce on children, relies heavily on research designs that create comparison groups of this type.

A major difficulty encountered in carrying out longitudinal designs is attrition, that is the loss of respondents between interviews due to death, mobility, refusal to continue in the study or some other cause. Like low response rates attrition may pose a problem for the generalizability of a study's results if the respondents who have attrited share a characteristic that seriously biases the representativeness of the sample. It is essential then, in evaluating research based on longitudinal designs, to carefully scrutinize attrition processes and the corrections that the researchers may or may not have made in their analyses to account for attrition.

Although longitudinal survey designs of the type we have described avoid many of the problems associated with the collection and use of retrospective data and although these designs greatly facilitate the researcher's ability to pursue comparative analyses, such efforts require the expenditure of significant sums of money. What is notable is that the research community, with the support of the Federal government and private foundations, increasingly has demonstrated a willingness to make precisely this sort of commitment.

Two examples of longitudinal national surveys that will allow for very sophisticated analyses of the American families and children are the National Survey of Children (NSC) and the

National Survey of Families and Households (NSFH).[44] The NSC began as a nationally representative survey of children who were first interviewed in 1976-77 when they were between the ages of seven and eleven. The same children were followed and reinterviewed in 1981 and again in 1987-88.[45] Importantly, information was also gathered from the children' parents or guardian and from their teachers. The purpose of the NSC was to provide the first highly detailed description of a representative cohort of American children as they moved from pre-adolescence into young adulthood. Beyond its basic research and descriptive goals, the NSC also focused on specific policy concerns in its second and third waves. Interviews in 1981 gathered extensive supplemental information on the consequences of marital conflict and divorce and interviews in 1987-88 focused on teenage sexual behavior and childbearing.

The NSFH is a large national survey of over thirteen thousand U.S. households and families that was fielded in 1987-88. The members of the original survey were reinterviewed in 1992-93 and the second set of information should be available to

---

[44] For general descriptions of these data sets, see respectively, Nicholas Zill, James L. Peterson, Kristin A. Moore and Frank F. Furstenberg, Jr. 1976-87 National Survey of Children: Waves 1,2, and 3 Child Trends, Inc. Los Altos, CA: Sociometrics Corporation (1992) and James Sweet, Larry Bumpass and Vaughn Call, The Design and Content of the National Survey of Families and Households (NSFH Working Paper (No.1)), Madison, WI: Center for Demography and Ecology, University of Wisconsin (1988).

[45] The National Survey of Children was used by Cherlin, et al., supra note 2; and by Frank Furstenberg, et al., The Life Course of Children of Divorce: Marital Disruption and Parental Contact, 48 Amer. Soc. Rev. 656 (1983).

BEST COPY AVAILABLE

the research community by early 1994. The 1987-88 survey interviewed adults, but the 1992-93 survey interviews both the original adults and a sample of children as well as adults who may have left the household due to events such as divorce.

The purpose of the NSFH was to develop a comprehensive national longitudinal data base on American families and the recent changes they have experienced. To gain some appreciation of the size of the NSFH it is notable that the files of certain family types in the 1987-88 sample alone can contain as many as two to three thousand pieces of information.[46] Interestingly, substantial oversampling of certain family types was intentionally undertaken to increase the number of cases available for detailed analyses of these family types. Oversampling refers to a practice used by researchers to assure that they have sufficient numbers of cases in their samples for analyses involving important events or individuals that occur infrequently in the population. As long as the oversampled cases have been selected randomly, and as long as the factor by which these cases have been oversampled is known, the sample can be weighted so as to remain representative of the population. The oversampled family types in the NSFH included single parent

---

[46] See Codebooks and Documentation: National Survey of Families and Households (Primary Respondent Questionnaire, Self-Administered Questionnaire, Spouse/Partner Questionnaire for Secondary Respondent). Center for Ecology and Demography, University of Wisconsin- Madison (1988). These estimates are based on our work with the NSFH.

BEST COPY AVAILABLE

families, cohabiting couples, and reconstituted families.[47] These are categories of families that do not fit the traditional married couple model that is the basis for so much of family law. Research on the NSFH families and households could contribute significantly to legal analysis and reform efforts, and the public availability of the NSFH represents an implicit invitation for legal scholars of the family to work cooperatively with social scientists to answer the questions they consider important for family law analysis and formation.

The NSC and the NSFH are very high quality data sets that have been and continue to be used for research on a wide variety of topics. They represent substantial scientific progress in the effort to understand American family systems, especially when considered in relation to prior non-longitudinal, non-probability convenience samples. In considering the merits of research based on probability and non-probability convenience sampling, it is important to take an historical perspective. Early in the development of research on a new topic, convenience samples will be more common and more acceptable because exploratory research is typically underfinanced, researchers are not sure what questions should be asked and measurement strategies are underdeveloped or nonexistent. But as research in an area develops, the social scientific community's standards for samples are adjusted upward and its tolerance of nonprobability samples

---

[47] Sweet et al , supra note 44, at 18. Minority families and recently married persons were also oversampled.

lessens, as they properly should. Given this historical context, it is natural to expect that the influence of findings from work based on convenience samples should lessen where it deviates from findings based on more rigorous research designs. The fate of each generation's scientific findings is to be superseded by the work of the next generation. This historical process is ongoing in the scientific community and it is both the great joy and the frustration of doing science.[48]

Hence, those wishing to use social science research in family law analysis and policy formation need to aware of the state of the field and to assess sample quality according to the maturity and sophistication that might reasonably be expected at the current stage of development of the research process.

C. Problems of Causal Inference

In this section we analyze how the social sciences approach the issue of assessing whether a causal relationship exists between an independent variable, i.e. a suspected cause, and a dependent variable, i.e. a suspected effect. In pursuing this exploration, we will develop a series of hypothetical and actual examples that illustrate both the difficulties involved in establishing causal relationships in family social science and several of the major strategies that have been developed to cope with these difficulties. In particular, we describe how

---

[48] See Max Weber, Science as a Vocation, in From Max Weber: Essays in Sociology 129 (H.H. Gerth and C. Wright Mills eds. 1958).

experiments, quasi-experiments and longitudinal surveys may be used in family research to evaluate policy-relevant causal relationships.

At a very basic level, the problem a family researcher faces in efforts to assess causal relationships is akin to the problem commonly faced by laboratory researchers seeking to evaluate the impacts of an experimental drug. It is with this analogy to the laboratory setting that we can begin our reflection on how the scientific method ideally approaches the work of making causal inferences.

In a clinical research setting, a researcher who is interested in the evaluating the impact of a drug typically will construct experimental and control groups of subjects. Experimental groups will receive the new drug in various regimes and the control groups will receive no drug, a placebo or the best drug currently in use for the relevant malady, the common cold for example. The researcher's goals are to compare the conditions of the experimental and control groups after each group has been exposed to the regime required by the design, and to be as certain as possible that no factor other than the drug treatment is causing any observed positive or negative impacts on the subjects. For example, the researcher would not want the experimental and control groups to significantly differ at the outset of the study regarding age, prior health condition or any other factor that could confound her understanding of the drug's impact. Thus the experimental and control groups should be

comparable in all respects, other than the treatment they received, so that the researcher will be able to unambiguously attribute any change in the subject's condition to the different drug treatments they have received.

This illustration leads one to ask if there is a way to construct the experimental and control groups such that they are in fact comparable on every factor that might confound the assessment of the impact of the drug. Scientists have concluded that the best way to equalize experimental and control groups is to use random assignment, that is chance and chance alone, to allocate subjects to experimental or control groups.[49] Random assignment, or randomization, does not absolutely assure that the experimental and control groups will always be equivalent (in flipping a fair coin, on occasion one will get seven heads in a row due to chance), rather it maximizes the long-term chances of achieving equivalence between groups and it allows the researcher to assess quantitatively the likelihood that equivalence has not been achieved.

Studies that employ random assignment to create experimental and control groups possess high levels of **internal validity**, that is by making their comparison groups very similar on all

---

[49] For the standard treatment of experimental design and randomization in the social sciences see Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (1966). For a valuable discussion of experimental research designs for public policy analysis see John Gilbert et al., Assessing Social Innovation: An Empirical Base for Policy in Statistics and Public Policy (W.B. Fairly & F. Mosteller eds. 1977).

BEST COPY AVAILABLE

characteristics that might influence the outcome of interest, even important characteristics unknown to the researcher, they control for alternative hypotheses and rival explanations of the drug's impact. Consequently, we can be more certain of the causal results of a study that uses a random assignment than one that does not.[50]

One of the problems with random assignment experimental research designs is that for ethical and political reasons they cannot always be carried out with human populations.[51] For example, randomly assigning custody between fit parents on divorce to test a primary caretaker presumption would raise political and ethical concerns. Or consider the issues involved in studying the impacts of innovations in a state's child support collection system. If researchers considered internal validity alone, a random assignment experiment in which some counties implemented the innovations and others did not would clearly be called for on scientific grounds. Yet even if there were no objections to this design on ethical grounds, state legislatures,

---

[50] Random assignment or randomization is commonly confused with random selection. Random selection is the process whereby we select a random sample from a population; its result is a probability sample. Random assignment or randomization is the process whereby we use chance to create an experimental group(s) and a control group(s). These processes are entirely independent research procedures.

[51] For a thorough discussion of design, implementation and ethical experimentation in the justice system, see Experimentation in the Law: Report of the Federal Judiciary Advisory Committee on Experimentation in the Law (Federal Judicial Center, 1981).

BEST COPY AVAILABLE

whose approval would be required to implement such a design, would typically be very reluctant to allow "chance" to determine which jurisdiction would receive the child support reforms. Hence, political concerns likely would bar the use of randomization.[52]

Given these limits, policy researchers often employ alternatives to pure experimental designs that do not involve random assignment, but retain as many internal validity strengths of randomized designs as possible, alternatives generally referred to as **quasi-experimental** designs.[53] One such alternative, used in the Wisconsin Child Support Assurance Study, involved producing matched pairs of counties in which only one county of each pair would implement certain child support innovations.[54] The counties were matched in that researchers, based on available information, paired counties on a limited number of important characteristics thought to influence child support behavior. Hence, in matching, comparison groups are

---

[52] There are instances of pure randomized experimental designs being used in family policy research but they are relatively rare. Perhaps the best known recent examples are the Seattle-Denver Income Maintenance Experiments that evaluated the impact of a negative income tax as an alternative to the current welfare system, Aid to Families with Dependent Children. See SRI International, Final Report of the Seattle/Denver Income Maintenance Experiment, Vol. 1 (Design and Results) (1983).

[53] On quasi-experimentation generally see Campbell & Stanley, supra note 47. On the use of quasi-experimental designs in policy and legal research see Donald T. Campbell, Reforms as Social Experiments, 24 Am. Psychologist 409 (1969).

[54] On the use of matched group designs in quasi-experimental research see Rossi & Freeman, supra note 9, at 219-27.

BEST COPY AVAILABLE

produced, based on the researcher's knowledge of the population and child support processes being investigated, rather than the more scientifically rigorous procedure of random assignment. Matching, as an alternative to a pure randomized design creates valuable comparison groups by which to assess the impact of the reforms so long as great care is taken in matching the groups. Thus, while the internal validity of a matched group design would not be as strong as that of a randomized experimental design, it could be nearly as strong and it will clearly be superior to a design that does not employ a control group.

The evaluation of the Wisconsin Child Support Assurance System employed this quasi-experimental matched group design in a demonstration study of the effect of immediate withholding of child support awards from the obligated parent's income. To do the pilot study, the state of Wisconsin contracted with ten counties to implement immediate wage withholding from all new child support obligors.[55] These ten counties were matched with ten similar control counties on a number of characteristics such as population size and composition, divorce rate, unemployment rate, and average per capita income, thereby creating comparison groups of counties with and without the child support innovations. Further, since the demonstration was to last three years (1983-1985), three years of predemonstration count data

---

[55] Garfinkel and Klawitter 1990, supra note 1. The Family Support Act of 1988 (Pub. L. 100-485) required states to have child support guidelines by October 1, 1989 and to use wage withholding by January 1, 1994.

(1980-1982) were also collected to allow for comparisons of collections before and after the implementation of automatic wage withholdings.[56] These data were supplemented further with surveys of judges, family court commissioners and other sources.[57]

A major advantage of this design is that because the comparison groups were matched on many important and potentially confounding variables, the researchers' ability to draw strong causal inferences about the impacts of the child support reforms was greatly enhanced, that is the study possesses high levels of internal validity. This type of design was feasible because the subject of the policy innovation was not highly controversial. There is general agreement that child support should be paid and that evaluating means to accomplish this goal effectively seems appropriate and not unfair. Further, the amount that obligors owe would not change, rather the collection mechanism was altered. In contrast, for family policy issues such as custody, it would neither be possible nor ethically defensible to create comparison groups either by random assignment or by some types of matched group designs.

Obviously the Wisconsin study involved significant effort and planning, including the need for special legislation. Because of the expense and difficulty involved in implementing experimental and quasi-experimental designs and using

---

[56] Garfinkel & Klawitter, supra note 1, at 159.

[57] Garfinkel 1986, supra note 1.

representative samples, many researchers rely on small, non-random samples and do not use control groups. Frequently such studies draw their convenience samples from clinical settings.[58] For example, as we noted in our discussion of sampling, Judith Wallerstein and Joan Kelly, in their study of the impact of divorce on children, used a sample of divorcing couples with children that had volunteered for short-term counseling, a sample likely to be highly unrepresentative of the general population of divorcing couples and their children.

But as important as this sampling problem was the fact that Wallerstein and Kelly had no control group, that is there was no group of parents and children similar to their subjects who had not experienced divorce with which to make comparisons. This design flaw limited the usefulness of their study even further. Without a control group, Wallerstein and Kelly were simply unable to substantiate scientifically that the problems of adjustment that the children in their sample encountered were the result of marital dissolution rather than some other process such as adjustment to adolescence that most children experience, independent of divorce. That is, without a well designed control group they could not rule out rival explanations to the one they proposed. In short their study had very low internal validity.

In addition to experimental and matched-group designs, a

---

[58] See Surviving the Breakup, supra note 12. For criticism of their methodology, see, Carol S. Bruch, Parenting at and After Divorce: A Search for New Models (Book Review), 79 Mich. L. Rev. 708 (1981).

third research design often used in family social science for studying causal relationships should be considered, namely the survey.[59] Surveys typically involve two major steps, the sampling of a population of interest, ideally using a probability sample, and the administration of some type of questionnaire or interview to members of the sample. In family research, social scientists increasingly conduct longitudinal surveys in which the same respondents are repeatedly queried over time. Examples of well designed longitudinal surveys already discussed are the National Survey of Children and the National Survey of Families and Households. Fortunately it is feasible to use survey data, especially if they are longitudinal, to both describe populations and to address causal issues relevant to family law analysis and formation.

To explain how this can be done, consider the use of data from a well designed longitudinal survey of divorced families to assess the potential impact of joint legal custody arrangements, relative to other custody arrangements, on post-divorce coparenting. Recall that the key to making causal assessments is the development of a control group that is initially equivalent to the experimental group that experiences the treatment or event that we wish to assess, in this case joint legal custody. With a longitudinal survey of divorce cases we can construct a control

---

[59] On the use of survey data for causal analysis see Travis Hirachi & Haman C. Selvin, Principles of Survey Analysis (1973), Rossi & Freeman, supra note 9, at 227-234 and James Davis, Elementary Survey Analysis (1971).

group of divorced families that have not experienced joint legal custody and a treatment group that has experienced joint legal custody. But how comparable are these groups? How much internal validity would conclusions concerning the impact of joint legal custody on children based on comparisons between these groups possess?

To answer these questions it is important to recognize that statistical techniques can be employed with survey data to control for, that is, make equivalent, background differences between groups we wish to compare that might confound our ability to make the sorts of comparisons needed for strong causal inferences.[60] This is precisely the approach taken by Maccoby and Mnookin in their study of post-divorce custody arrangements in two, highly diverse California counties. In this study a longitudinal survey design was employed in which a court caseload of divorces was sampled soon after initial filings and the divorcing parents were interviewed on three occasions during the three years after the divorce. One of the issues that Maccoby and Mnookin explored in these analyses was the effect of joint legal custody arrangements on post-divorce parenting, visitation and child support. In making comparisons between joint-legal custody and other custodial arrangements, the authors developed

---

[60] For discussions of the value of longitudinal data in assessing causal relationships see James S. Coleman, Longitudinal Data Analysis (1981); Jay Teachman, Methodological Issues in the Analysis of Family Formation and Dissolution, 44 J. of Marriage and the Family 1037 (1982) and Paul D. Allison, Event History Analysis: Regression for Longitudinal Event Data (1984).

BEST COPY AVAILABLE

statistical models that controlled for potentially confounding background factors such as pre-divorce parental income and education. For example, in controlling for pre-divorce family income, the authors were able to account for the possibility that higher income families may be both more likely to choose joint-legal custody and more likely to experience or report better post-divorce adjustment. In fact, Maccoby and Mnookin found that joint-legal custody had little effect on post-divorce parenting, visitation and child support.[61] These findings are credible in terms of internal validity because of the rigorously developed statistical models that the authors developed to control for differences between divorced families with joint-legal custody and those with other custody arrangements. It is important to be aware that statistical controls of the type used by Maccoby and Mnookin would not be considered as robust as the controls produced by a randomized or a quasi-experimental matched-group design, but that comparisons derived from powerful statistical models do provide reasonably strong grounds upon which to draw causal inferences.

The important lesson to be drawn from this discussion is that in using social science research in family law analysis and formation, it is essential to ask the internal validity question and to examine critically the grounds upon which researchers make causal assertions. All social science research designs are not

---

[61] Maccoby & Mnookin, supra note 4, at 175 (no effect on child support), 225 (no effect on parental cooperation in decision making), 257 (no effect on visitation).

equal in their ability to control for alternative explanations and rival hypotheses, and the proper use of social science findings requires disciplined inquiry concerning questions of causal inference.

D. Researcher Bias

In addition to problems in measurement, sampling, and causal inference, researcher bias is another cause for concern in using social science in family law analysis and formation. By bias we mean the inadvertent influence of both the overt expectations the researcher may have had about the outcome of the research, and the covert beliefs of the researcher that may affect the research without the researcher anticipating their influence.[62]

First, consider the bias coming from the researchers' acknowledged expectations. A research project is designed to test overt theoretical expectations or hypotheses about empirical reality, but researchers' findings often do not support these hypotheses, at least not entirely. Researchers trained in the scientific method are taught to expect unanticipated findings and incorporate these new findings into the ongoing process of theory building. In this sense, the research process can be described as having a strong tendency toward "objectivity." The researcher poses a hypotheses or states a prediction, tests it against empirical reality, and is willing to accept the answer, even if

---

[62] We do not include outright dishonesty as a category of bias. Although of research fraud is a serious problem, in this paper we address the problem of inadvertent influence only. On "fudging" research results, see, e.g., Research Craft, supra note 9, at 96-97.

BEST COPY AVAILABLE

it was not the anticipated result. Since researchers know what they anticipate, they are expected to ensure that their research is not slanted so as to favor the anticipated answer. The norm of the scientific community is impartiality, that is openness to the possibility that one's hypothesis may not be supported by empirical analysis, rather than advocacy. In contrast to the scientist, the legal advocate has a goal, and structures and interprets the facts and law to achieve that goal. One question related to bias, therefore, is the extent to which scientists successfully avoided slanting their research to favor their expectations.

A second, more insidious bias problem comes from the ways in which the unidentified beliefs, prejudices or assumptions of the researcher may affect the research. These influences are harder, and probably impossible, for the researcher or a critic to entirely identify and eliminate.

Related to these broad considerations about impartiality in scientific research are recent complaints about bias in the use of social science in family law analysis and formation that have identified two concerns: first, the way the research is presented, namely that unwarranted claims of objectivity are made; and second, the existence and nature of bias itself.[63] First, consider the complaint that researchers claim to be objective.

1. Claims of Objectivity.

---

[63] See Fineman & Opie, supra note 5.

The basis for this complaint is that researchers are either unaware, or fail to acknowledge, their biases, and that they claim their research is value-free, thereby giving their research a gloss of credibility that is inappropriate.[64] This criticism is not entirely accurate. Current social science methodologists openly acknowledge the problems of bias and would suggest that anyone who thought the research was value-free was simply uninformed.[65] Indeed the 1993 Annual Meeting of the National Council on Family Relations will address the question:"What are the ethical and values assumptions upon which we base our study of families and our work for families?". The agenda for the meeting is directed at exploring bias.

The scientist's goal is to strive toward unbiased work, but the view that totally value-free work will actually be achieved has been criticized as scientifically naive for some time.[66] Part of the function of research methodology reporting norms is to make clear to the reader how the research was conducted so that the reader will be able to make an informed judgement about the quality of the research, including problems of bias. Hence, it is common practice for researchers to describe, in survey

---

[64] Id.

[65] The subject of researcher bias is addressed in standard, introductory texts on social science research. See, e.g., Research Craft, supra note 9, and Babbie, supra note 9, at 76, 88-90.

[66] See, e.g., Thomas Kuhn, The Structure of Scientific Revolutions (2d ed., 1970) and Faigman, supra note 5. Some feminists have claimed that feminist researchers can eliminate sexist bias in their research, but this assertion also has been criticized. See infra text accompanying notes 76-78.

BEST COPY AVAILABLE

research for example, the way the data were collected and analyzed, and to make copies of the data collection instruments available to the research community. Indeed, for much federally funded research, the data must be archived so that it is available to other researchers to reanalyze. Social scientists operate on the assumption that others will criticize their work and that their scholarly work will undergo double-blind peer review before publication in professional journals. After publication, scientists expect that others will seek to replicate their findings. These processes of accessibility, criticism, and replication are intended to create incentives for improving research generally, by directly addressing problems such as bias. This system is far from ideal. Researchers are heavily invested in their work and bias clearly influences the research process. Further, institutional procedures for reducing bias are not uniformly effective. The innovation of the scientific method is not that it has or ever will eliminate bias, but rather that it has created a socialization process and a system of institutional norms and procedures that address the problem of bias as a integral part of doing "everyday science."

The complaint is accurate, however, that scientific writing, like much professional writing, traditionally uses an objective tone that may be misleading. In writing about the natural sciences, Professor Ruth Hubbard notes that when she reports a discovery she does not write in the first person but rather uses

phrases such as "'It has been observed that . . ..'"[67] In criticizing this practice she notes that:

> This removes relevance of time and place, and implies that the observation did not originate in the head of a human observer, specifically my head, but out there in the world. By deleting the scientist-agent as well as her or his participation as observer, people are left with the concept of science as a thing in itself, that truly reflects nature. . . .[68]

So although scientists may not believe that their work is value-free, they may present it as value-free, and typically do not include explicit consumer warning labels about bias. Such warnings are left implicit for they are taken for granted by the scientific community. Professionals in other fields should not be deceived by an objective tone, however. Law reviews, for example, probably do not need to indicate that objective statements by an author such as "the better approach is that" or "the correct interpretation is" merely represent the author's opinion, not objective truth. When writing for a lay audience, however, researchers should be careful to indicate the limitations of their work. Unfortunately some family researchers, when writing books for a lay audience, have made stronger claims for the general validity of their findings than

---

[67] Ruth Hubbard, Science, Facts, and Feminism, in Feminism and Science 119, 126 (N. Tuana, ed. 1989).

[68] Id.

A major difficulty with attempting to explain feminist critiques of bias in social science research succinctly is that there is no one feminist position. Rather, there are many criticisms from different feminist perspectives and although they consistently assert the existence of bias, they do not agree on how and whether it can be corrected. Professor Sandra Harding has identified three main themes or epistemologies among feminist critics of social science. These critiques are made by the feminist empiricists, the standpoint theorists, and the feminist post modernists.[71] A fourth type of feminist criticism, which we refer to as an anecdotal approach to social science, was used by law professor Martha Fineman and her colleague Anne Opie.[72] Their criticism is more narrowly tied to the use of social science in policymaking rather than being a general epistemology.

Of course, not all feminist criticisms fit neatly into these four categories and as feminist criticism evolves, expanded or new definitions of categories of feminist criticism are necessary. An additional category that has been suggested, for

---

empiricism and sociolegal studies in 14 L. and Soc. Inquiry (1989).

[71] Sandra Harding, The Science Question in Feminism (1986) [hereinafter Harding (1986)]; Sandra Harding, Conclusion: Epistemological Questions, in Feminism and Methodology 182 (S. Harding, ed. 1987) [hereinafter Harding (1987)]; and Sandra Harding, Is There a Feminist Method, in Feminism and Science 17, 25 (1989).

[72] Fineman & Opie, supra note 5.

BEST COPY AVAILABLE

example, is the postempiricists,[73] but we have included a discussion of this category in an expanded review of the empiricists category. For purposes of this Article, maintaining the four categories is useful to demonstrate a variety of concerns about bias. Taken as a whole, the feminist critics show that bias is pervasive and unavoidable.

The first category of critics of social science, the feminist empiricists, generally believe that sexist and androcentric bias in mainstream social science research can be reduced or eliminated by stricter application and adherence to standard methodological rules for conducting research.[74] These feminists believe that research that has its origins in feminism can be more accurate than research that attempts to be neutral, a paradox noted by Harding.[75] The early feminist empiricists believed that bias could be eliminated.[76] They argued that feminist researchers (both male and female), who will have fewer

---

[73] See Joyce McCarl Nielson, Introduction in Feminist Research Methods: Exemplary Readings in the Social Sciences 1, 27-31 (Joyce M. Nielson, ed. 1990) [hereinafter Feminist Research].

[74] Harding (1987), supra note 71, at 182-83.

[75] Id. at 181-82.

[76] Professor Linda Alcoff suggests that this claim has been made in the following works by feminist scholars: Margaret Eichler, The Double Standard: A Feminist Critique of Feminist Social Science 51 (1980); Evelyn Reed, Sexism and Science 7 (1978); Nancy S. Dye, Clio's American Daughters: Male History, Female Reality, in The Prism of Sex: Essays in the Sociology of Knowledge 22, (1979); and Ruth Hubbard, Have Only Men Evolved, in Discovering Reality 45, (1983). Linda Alcoff, Justifying Feminist Social Science, in Feminism and Science 85, 88 (1989).

BEST COPY AVAILABLE

person asks questions, and the other answers. The interviewees are passive; the interviewers have a question asking, rapport building role.[80] The goal of this type of interview is to have all the interviews be comparable and to avoid influencing the respondent thereby enhancing measurement validity. This form "appeals to values such as objectivity, detachment, hierarchy and 'science'..."[81]

Oakley explained how this model of interviewing would have been inappropriate and unsuccessful in research she conducted on the transition to motherhood. In her study, fifty-five women were interviewed four times-- twice during pregnancy and twice afterwards, with an average total interviewing time of 9.4 hours.[82] She even attended the birth of the babies of some of the women in her study. She explained that she was asking a great deal of these women; not just their time, cooperation and hospitality at a time when they might want to exclude strangers, but also "confidences on highly personal matters such as sex and money and 'real' (i.e. possibly negative or ambivalent) feelings about babies, husbands, etc...."[83] She felt that the use of the prescribed interview would have been morally indefensible. Rather, she felt that she needed to respond to the women's questions on factual matters relating to childbirth, about her

---

[80] Id. at 34-36.

[81] Id. at 38.

[82] Id. at 41.

[83] Id. at 43.

problems, develop hypotheses, and test hypotheses in ways that are different from mainstream researchers.[87] The concept of a feminist methodology is still evolving, however. Although many discussions of feminist methodology focus on the use of qualitative methods, the discussion has broadened to include quantitative methods as well and to raise basic questions about methods of inquiry generally.[88] In reviewing feminist research in sociology, for example, five feminist epistemological principles have been identified. These include:

> (1) the necessity of continuously and reflexively attending to the significance of gender and gender asymmetry as a basic feature of all social life, including the conduct of research; (2) the centrality of consciousness-raising as a specific methodological tool and as a general orientation or "way of seeing"; (3) the need to challenge the norm of objectivity that assumes that the subject and object of research can be separated from one another and that personal and/or grounded experiences are unscientific; (4) concern for the ethical implications of feminist research and recognition of the exploitation of women as objects of

---

[87] For a discussion of feminist methodologies, see Judith A. Cook and Mary Margaret Fonow, Knowledge and Women's Interests: Issues of Epistemology and Methodology in Feminist Sociological Research in Feminist Research, supra note 73, at 69.

[88] DuBois, supra note 78, at 109; Toby Jayaratne, The Value of Quantitative Methodology for Feminist Research in Theories 140, supra note 78; Maria Mies, Towards a Methodology for Feminist Research in Theories, supra note 78.

BEST COPY AVAILABLE

knowledge; and (5) emphasis on the empowerment of women and transformation of patriarchal social institutions through research.[89]

The evolution of the feminist empiricist critique should continue to provide valuable lessons for the design and use of social science research. Of the four categories, this critique most readily allows social science to continue as a "science" that both men and women can participate in with the common goals of reduction of bias and improved understanding of society.

In addition to the feminist empiricists, a second category of feminist critics of bias in social science is the standpoint theorists.[90] The feminist standpoint theorist takes the position that the feminist social experience is more complete and less distorted than the dominant male perspective. Feminist research, therefore, can be more accurate because it is developed and tested against this more comprehensive social experience.[91] For some, this would mean that the research would need to be done by, on and for women. This would result in a feminist social

---

[89] Cook & Fonow, supra note 87, at 72-73.

[90] Harding (1987), supra note 69, at 184. According to Harding, some influential papers on the epistemology are: Jane Flax, Philosophy and the Patriarchal Unconscious: A Psychoanalysic Perspective on Epistemology and Metaphysics, in Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology and Philosophy of Science ___ (Harding and Merrill Hintikka, eds. 1983); and Hilary Rose, Hand, Brain and Heart: A Feminist Epistemology for the Natural Sciences, 9 Signs: J. of Women in Culture and Soc. ___ (1983). Harding (1987), supra note 71, at 6. Harding cites other sources in this note as well.

[91] Id. at 184-85.

science that was initiated out of women's experience of women's reality. Without this we would only have a "social science in which women's experiences are researched and analyzed using the conceptual procedures, the methods of research, and the research models provided by sexism."[92] Some feminists have suggested that we will not have a feminist science until we have a feminist society.[93] Others have attempted to describe what feminist science might do. An account of housework, for example, from the standpoint of women might differ substantially from an account by men. Time-budget studies have classified housework as part labor (similar to wage work for others) and part leisure (self-directed activity). Sociologist Dorothy Smith suggests, however, that:

> [F]or wives and mothers, housework is neither wage labor nor self-directed activity. An account of housework from "the standpoint of women" --our experience of our lives--rather than in terms of masculine science would be a quite different account; the voice of the subject of the inquiry and the voice of the inquirer would be culturally identifiable.[94]

---

[92] Liz Stanley & Sue Wise, Breaking Out: Feminist Consciousness and Feminist Research 165 (1983). (Alcoff suggests that Stanley and Wise present standpoint theorists claims. Alcoff, supra note 76, at 89.)

[93] Harding, 1986 supra note 71 citing Elizabeth Fee, Women's Nature and Scientific Objectivity in Women's Nature: Rationalizations of Inequality (M. Lowe & R. Hubbard, eds. 1981).

[94] Harding, 1986 supra note 71 citing Dorothy Smith, A Sociology for Women in The Prism of Sex: Essays in the Sociology of Knowledge (J. Sherman & E.T. Beck, eds. 1979) and "The Experienced World as Problematic: A Feminist Method," Sorokin lecture no. 12,

The feminist standpoint theorists have been criticized, however, for approaching radical relativism. Professor Linda Alcoff succinctly states this criticism:

> Uncareful articulations of value-laden theory-choice could lead to the view that the differences between feminist and androcentric social scientific theories are merely political differences, which do not render the theories true and false but only different. In other words, it might lead to a conception of social scientific theories as having an empirical content and a value content, autonomous from each other in such a way that two conflicting theories may share their empirical content but contradict each other's value content. In such a scenario, the debate over theory-choice would necessarily recede to the political and/or ethical domain entirely.[95]

Alcoff is concerned that this situation would undermine the ability of feminist social scientists to provide empirical support for feminist claims.[96]

A third category of feminist critics are the feminist post-modernists, who question the value of any science, "feminist" or otherwise.[97] This skepticism comes from concern that it simply

---

Saskatoon, University of Saskatchewan (1981).

[95] Alcoff, supra note 76, at 89.

[96] Id. at 90.

[97] Harding (1987), supra note 71, at 187-88. Harding recommends the following for discussion of postmodernist issues: Jane Flax, Gender as a Social Problem: In and For Feminist Theory

BEST COPY AVAILABLE

is not possible to accurately describe a part of the world or reality as a unitary entity or phenomenon, and that attempts to do so ignore the complexity, the many stories, and the different voices that, in fact, make up reality. These feminists are opposed to attempts at a dominant, unitary story and hence are opposed to any science, even "feminist" science.

This argument does not address the fact that our concept of the "unified voice" is also part of our reality. We do group people that we perceive as being alike into a collective and ascribe to them a particular point of view and these perceptions are used in policymaking. One of the advantages of social science is that it can provide a disciplined conceptual framework for identifying that ascribed voice and challenging it by giving the unheard a way to speak, albeit artificially and collectively. A stereotyped view of recipients of Aid to Families with Dependent Children (AFDC), for example, is that they are unwilling to work. Social science can require us to define and measure "unwilling to work" and allow us to question whether this is a typical characteristic of AFDC recipients relative to other populations. Although a study of recipients that produced a "group answer" would obscure the individual responses, the group answer, the unified voice, has value as well.

Further, collective experience was one of the strengths that came from consciousness raising, because women discovered that

---

in American Studies/Amerika Studien (1986); New French Feminisms (E. Marks and J. Courtivron, eds. 1980); and essays in Signs: J. of Women in Culture and Society (ed. 1981).

they shared many common experiences and thoughts. This provides both a way of confirming experiences of women that have been ignored and of developing a new sense of reality and self. The collective knowledge can have value both for individual self-knowledge and group action. The feminist empiricists advocate the use of consciousness raising as a methodological tool. Hence, a unified story can be useful to feminist work, even though it is not an accurate portrayal of the experience of any particular individual, or even of groups of individuals within the larger category.[98] Although we need to be sensitive to the multiple voices within a category, we also need the category.[99]

A fourth type of criticism, an anecdotal approach to social science, was used by Fineman and Opie in an article on child custody and has been developed further by Fineman in subsequent work.[100] Unlike the other critiques, the anecdotal approach was developed in the context of reviewing existing research for the purpose of policymaking, rather than focusing on the conduct of

---

[98] See Marcia Westkott, Feminist Criticism of the Social Sciences, in Feminist Research, supra note 73 and Cook & Fonow, supra note 87 at 75.

[99] In the context of feminist legal theory, Prof. Angela Harris has argued that "[a]bandoning mental categories completely would leave us terrorized by the sheer weight and particularity of experience. No categories at all, moreover, would leave nothing of a women's movement, save perhaps a tepid kind of I've got my oppression, you've got yours' approach." She goes on to emphasize the need for awareness of diversity and multiple consciousness, however. Angela Harris, Race and Essentialism in Feminist Legal Theory, 42 Stan. L. Rev. 581 (1990).

[100] Fineman & Opie; Fineman,1989; and Fineman, 1990 supra note 5.

BEST COPY AVAILABLE

research. We have included it here rather than in the section on the research review, however, because concerns about using research done by others seem closely related to critiques of doing research. The Fineman and Opie approach judges research by its results in relation to their vision of women's interests, rather than by scientific or other epistemological criteria. This approach uses social science research like anecdotes or cases that can be employed to support a particular feminist political position. Since both the research and the selection process is value-laden, studies can be used selectively without the need to justify selection or omission in relation to scientific community norms governing research reviews.[101] In a review of custody research, for example, Fineman and Opie effectively took the position that research on the parenting ability of fathers that showed them in a favorable light could and should be omitted because it did not support the ideological position that mothers are the best caretakers for children.[102] Research, however, that favored mothers was included without regard to its methodological flaws.

---

[101] To some extent, Faigman suggests that the Supreme Court has followed a similar approach in using empirical research. The Court has used empirical research that supported a desired outcome and ignored research that did not. He argues that empirical research nonetheless serves to restrain the Court's decisionmaking. Faigman, supra note 5. For a discussion of the research review, see section III.B. infra.

[102] There is, of course, more than one feminist position on issues relating to mothers and custody. See e.g. Katherine T. Bartlett & Carol B. Stack, Joint Custody, Feminism, and the Dependency Dilemma, 2 Berkeley Women's L. J. 9 (1986) and text accompanying notes 115-116 infra.

56 BEST COPY AVAILABLE

Fineman is critical of other legal scholars who pretend to be objective when they are using social science to justify an ideological position.[103] She suggests that the training and culture of lawyers reinforces their tendency to be non-objective advocates.[104] She feels that the problem of selective incorporation of social science research by legal scholars is inevitable, but that social science should continue to be used. What should cease is the pretense of objectivity.[105]

Unfortunately, Fineman and Opie do not always follow this advice and appear at times to be attempting to review social science research objectively. They do state their ideological position, but they also use social science studies to support their position and criticize opposing positions without sufficiently emphasizing that the research is being used anecdoctally.[106] They also use the methodological flaws in undesirable research to justify why it is inappropriate, even though they use work with similar flaws. This attention to methodology lulls the reader into thinking that the research was

---

[103] Fineman,1989, supra note 5 at 107.

[104] Fineman, 1989, supra note 5 at 105.

[105] Id. at 107; see also Fineman, 1990 supra note 5 at 186.

[106] Their ideological position is that "in the vast majority of cases mothers present the best alternative for child custody after divorce" (at 111)). For their use of social science research to support their position and criticize opposing positions see e.g. their footnotes 20, 86, and 91. Fineman & Opie, supra note 5. Fineman's selective use of social science research in her book The Illusion of Equality: The Rhetoric and Reality of Divorce Reform (1990) is criticized in Milton C. Regan, Divorce Reform and the Legacy of Gender, 90 Mich. L. Rev. 1453, 1475-76 (1992).

BEST COPY AVAILABLE

57

selected based on scientific rather than an ideological criteria.

The anecdotal approach to social science uses ideology to select research. By excluding some types of research entirely, it differs from the feminist empiricists who propose that their research, based in an ideology, will be more inclusive than prior research that attempted to be politically neutral. The empiricists argue that by being more inclusive, their research will be less biased, and will therefore be better than prior research. They are concerned with improving the research process, rather than controlling the result. The Fineman and Opie approach, in contrast, is exclusive, not inclusive and is focused on the use of research results independent of the process that produced them.

Although Fineman and Opie are similar to the standpoint theorists who want research that is done for women, they differ from them in that they do not require that the research be conducted by women or be initiated out of women's experience. Since their focus is on research results, the researcher and research process are not controlling factors in judging research.

Unlike the post modernists, however, the anecdotalists want to use research that supports their position, rather than rejecting it entirely. The choice of research is directed by an ideological position rather than by scientific criteria. To the extent that they want the studies that support their position accepted as science, however, their approach, unfortunately, contains the radical relativism trap that Alcoff identified in

relation to standpoint theorists; namely that it would not allow feminist social scientists to argue that there was scientifically respected empirical support for feminist claims. Overall the anecdotalists' work is valuable for its emphasis on the need for concern about the lack of objectivity in the conduct and use of research. The major difficulty with their position is that they do not adequately explain why we should accept the research studies they cite to support their position, while rejecting or ignoring others.

These four feminist criticisms of social science research demonstrate the existence and pervasiveness of bias. Each of these four epistemologies identifies different concerns and raises new questions. The latter two in effect say that social science research has no real content and therefore should not be used at all (the post modernists) or should be used only if it assists an ideology (the anecdotalists). The standpoint theorists and the anecdotalists suffer from the radical relativism trap. Hence, we side with the empiricists because we believe social science research can be a valuable and dynamic voice in legal analysis and law formation. This leaves us with the problem of what to do about bias.

3. What Can Be Done about Bias?

Before taking up the issue of how best to deal with the problem of bias in family research, it will be useful to summarize two major themes that have emerged from our analysis of

objectivity and bias up to this point. First, we conceptualize objectivity as a goal that should be sought, but which in principle is not achievable. Second, objectivity is a process that is historically and socially grounded both in the scientific community's many failures to deal with bias that has distorted research and its use, and its dedication to institutionalized and ongoing methodological criticism based in the scientific method as a means for dealing with bias.[107] Given this context, today's family social science, while certainly biased, is in all likelihood more objective than the family social science produced at the turn of the twentieth century because of the ongoing critical assessment by the scientific community. Improvements in family research can be seen in the higher quality measurement, samples and methods of causal analysis that characterize the best of family social science today, and, perhaps more importantly, in the progressive and inclusive manner in which methodological criticism of family social science developed out of perspectives such as feminism and multiculturalism have been incorporated by the institutional process of criticism that is science.[108]

---

[107] For an accessible treatment of the historical dimensions of "institutionalized criticism" and its role in the scientific method see Stephen J. Gould, The Mismeasure of Man (1981). Gould's analysis of 19th and 20th century racism and its influence on scientific measurement is particularly revealing in that it supports the dual notions that (a) science is an inherently human and therefore biased/value laden endeavor and (b) science is always correcting for its biases even if, in some instances, it takes a century to do so.

[108] For an important illustration of a social science, sociology, seeking to seriously respond to feminist criticism by developing guidelines for reducing gender bias in research

With this conceptualization of objectivity and bias as background, we turn to the questions of how to reduce bias in research and how to use research findings that are influenced by the researcher's values. One suggestion has been that researchers should reveal their biases,[109] so that the consumer of the research could then correct for them in some way. What biases should be revealed to achieve this goal is not self evident, however. Presumably the biases that the researcher is unaware of are the most dangerous.[110] The researcher could reveal some general information about her profession, background and family structure, for example, and the user of the research could then apply whatever stereotype about values those demographics conjure up and assess the research accordingly. Or the researcher could reveal her personal and political beliefs about each hypothesis in a long addendum-- her known biases. Although these efforts might help the researcher's self-examination for bias, it does not appear likely that either approach would significantly help others in assessing the research. Discussion of the researcher's values or life experiences would be helpful, however, when the researcher can

---

methodologies see, Committee on the Status of Women in Sociology, The Treatment of Gender in Research (1986).

[109] Faigman, supra note 5, at 1030; Fineman & Opie, supra note 5, at 125.

[110] See Faigman, supra note 5, at 1030.

identify and explain their influence.[111]

Another way to reduce bias is through replication. Replication refers to the scientific ideal of having findings corroborated repeatedly in independently conducted, but similarly designed, studies before they are generally accepted. Replication is quite common in laboratory sciences such as pharmacology and psychology, but more difficult to attain in sciences such as ecology, geology, sociology, demography and economics that employ field data rather than data collected in laboratory settings. The difficulty experienced by these disciplines, and by family social science in particular, in achieving replication to the same degree as laboratory sciences has several causes. First, large scale field studies such as national surveys are extremely expensive when considered in relation to the costs involved in conducting and replicating most laboratory studies. As a consequence, cost typically precludes the simultaneous or immediate fielding of replication studies in the social sciences. Second, even if funding was available to conduct multiple studies simultaneously in different locations, such studies would fall short of the laboratory ideal of replication because locations will almost surely vary to a much greater extent than laboratory settings. Finally, for similarly

---

[111] An interesting example of personal experiences shaping theory building can be found in Robin Paul Malloy, Is Law and Economics Moral?-- Humanistic Economics and a Classical Liberal Critique of Posner's Economic Analysis, 24 Val. U. L. Rev. 147 (1990). Professor Malloy explains how his theory development was related to his childhood experiences in an abusive home. Id. at 155-56.

62

BEST COPY AVAILABLE

designed studies not fielded at the same time, there would still be the problem of whether subsequent studies had actually been conducted in settings comparable to the original study's because of the passage of time and the influence of historical events. These difficulties should not lead one to conclude that replication never occurs in the social sciences, but rather that the process of replication is more difficult and more time consuming than in other fields.

Another way to reduce bias is by careful review of the research before and after publication by a community of scientists, many of whom increasingly work out of feminist and other critical perspectives and come from diverse ethnic and cultural backgrounds. Research can be critically reviewed for bias by asking the same kinds of questions that should be asked in assessing the quality of research generally,[112] such as: what was the conceptual framework of the study? What methodology was used and how appropriate was it? What was not asked--where are the silences? Are there other plausible interpretations of the findings? Are the conclusions consistent with the findings and with other research?

Critics of research clearly will be informed and motivated by their own bias, perspective or voice (the paradox that Harding notes of a political agenda resulting in less biased research).[113] A feminist critique of a design of a study on

---

[112] See Part II and infra text accompanying notes 135-37.

[113] See supra discussion accompanying note 75.

BEST COPY AVAILABLE

child custody decisions at divorce, for example, might question whether using a primary caretaker presumption as a major policy variable was useful or beneficial to women. Professor Herma Hill Kay has argued that we should encourage nonsexist, shared parenting, rather than single, primary caretaking.[114] She stated that the predictability that a primary caretaker presumption "brings to custody awards is purchased at the cost of legitimating the maternal preference under an easily penetrated veneer of gender neutrality that effectively excludes the vast majority of fathers as potential custodians."[115] She suggests that "the normative effect of such a legal preference actually might tend to discourage fathers from participating in the care of their children during marriage while reinforcing the existing cultural directive that women ought to regard mothering as their primary role." [116]

A feminist critique could also focus on methodological problems. In a study of the effects of custodial arrangements of children's well-being, for example, a critique could focus on the difficulties involved in successfully controlling for situational influences other than custody that may influence well-being. Since physical custody is so often with mothers, fathers with

---

[114] Herma Hill Kay, Beyond No Fault, in Divorce Reform at the Crossroads 6 (Sugerman & Kay, eds. 1990). Kay's criticism is just one example of a feminist critique. Other feminists, such as Fineman, favor the primary caretaker presumption.

[115] Id. at 35.

[116] Id.

custody are by definition unusual.[117] With such a limited number of cases, could we realistically hope to control for all the factors that would be necessary to feel confident that we were looking at the effect of the custody arrangement rather than something else? Would it not be necessary to explore how that custody arrangement came about? Do we know how to adequately control for the problems of poverty when assessing custodial outcomes? If not, this could adversely affect mothers who typically have a lower income than fathers.[118] Also, if both parents are important to the child, should we also consider the effect of the custody arrangement on the parents? Many feminists and others would also argue that a study of this complexity should include qualitative, longitudinal data for a substantial number of cases.

Finally, in addition to replication and peer review, we should continue to expect the researcher to be concerned about bias. Researchers need to be continually questioning their own honesty, fairness and motives in their research. Although this may cause doubts and uncertainties for the researcher it is a valuable and expected part of the scientific process. According to one feminist "I'd be surprised if there is anyone who is doing feminist scholarship who hasn't at some point wondered whether she were being honest and 'objective' in her work, whether she

---

[117] Andrew J. Cherlin, Marriage, Divorce, Remarriage 26 (Rev'd ed. 1992).

[118] Id. at 73.

weren't perhaps 'making it all up,' whether she were being manipulative or biased in her design, her reporting, her interpretations."[119]

Bias then can be reduced, although not eliminated, by a standard critique based on scientific principles, a critique based in another ideology, and by the conscious efforts of the researcher. This approach would help us identify the nature and extent of bias in the research, continue the process of reducing bias, and increase our understanding of the relationship between the knower and the known.

III. Using Research in Family Law Analysis and Formation

As our general description of basic problems of measurement, sampling, causal inference and bias has shown, methodological problems in family research are formidable and demonstrate the need for caution in using social science research in policy formation. The next Part emphasizes that empirical research is an on-going process that can inform a policy debate. Research does not, however, provide the answer to a policy question, because a policy choice is a normative decision. This Part also considers the research review, a developing methodology for analyzing and summarizing research for law reform and other purposes.

---

[119] Barbara DuBois, supra note 78, at 113.

A. The Need for Multiple Studies and the Normative Basis of Policymaking

The methodological discussion in the previous Part has primarily focused on study design and data collection. Of course, a number of judgments need to be made in the process of analyzing the data and interpreting the findings as well. A study may have collected data on a number of important independent and dependent variables, but the analysis may show that relationships among some of these variables were not statistically significant and that very little of the variation in the dependent variables of interest was explained, or that relationships that were found contradicted expectations. Part of the researcher's job is to study these results and explain whether or not they support the study's initial hypotheses.

Because of these difficulties, beginning studies in a complex area would be very likely to produce, at best, very tentative conclusions and provide very qualified policy guidance. Indeed it is highly unlikely that any one study could provide a complete picture of a complex issue.[120] Instead policymakers should consider a number of studies and use them to gradually increase the knowledge base concerning the policy issue and possible resolutions.[121]

---

[120] Richard J. Light & David B. Pillemer, Summing Up: The Science of Reviewing Research 159 (1984).

[121] This approach not only avoids overemphasizing one study inappropriately, but also avoids discounting the cumulative value of the studies because each has some compromises and flaws. Id. See also Phoebe C. Ellsworth, To Tell What we Know or Wait for

67 BEST COPY AVAILABLE

Even if we had a number of high quality studies with findings that were unambiguous and consistent, they still would not tell us whether to adopt the policy, although our decision presumably would be more informed than if we did not have these studies. No study or combination of studies will provide "the" answer.

To illustrate this point, consider the evaluation discussed earlier of the Wisconsin Child Support Assurance System. One of the goals of that evaluation was to determine the effect of wage withholding on child support collections. As is usual with policy research, the researchers had to deal with a variety of real world problems in carrying out this field study. A major problem was that the pilot counties did not fully implement immediate wage withholding and the control counties unexpectedly increased their use of immediate wage withholding during the study.[122] If the study had only comparisons between control and pilot counties, therefore, without taking these factors into account, the true effect of a fully implemented immediate wage withholding would have been underestimated because of the underutilization of wage withholding by the pilot counties and the increase in its use by the control counties.[123]

On the other hand, if the study had focused on the differences in child support payments between cases with and

---

Godot?, 15 L. and Hum. Behav. 77, 78-79 (1991).

[122] Garfinkel & Klawitter, supra note 1, at 161-65.

[123] Id.

without immediate wage assignments, an overestimation of the effect probably would have resulted because obligors who were unemployed or self-employed, who typically might be required to pay less because of lower income, were not subject to wage withholding. [124] Unfortunately the researchers were not able to adequately control for this possible overestimation.[125]

The researchers dealt with these problems by reporting more than one estimate. A comparison of the pilot counties to the control counties showed an 11% increase in child support payments due to immediate wage withholding. But a comparison of the child support payments for cases with wage withholding to payments for cases without wage withholding showed a difference of 30%.[126]

Both estimates showed an increase, albeit of different magnitudes, and these findings were consistent with other research, but note that these findings do not answer the question of whether the policy should be implemented statewide. Rather, they provide information with which to consider whether the policy should be implemented. Although using immediate wage withholding had a substantial positive effect on payments, the effect was small compared to other estimates of the payer's ability to pay child support, which would suggest that payments

---

[124] Id. at 174.

[125] Id. at 161.

[126] Id. at 174.

could be increased up to 400%.[127] A legislature might decide that attention should be focused on obtaining more child support through mechanisms such as paternity establishment programs and increasing award levels rather than diverting attention by focusing on the collection of existing awards. Others might be concerned with the unintended effects of wage withholding on payors. With low income payors, for example, would wage withholding, because it substantially increases the effective tax rate on earned income, have an adverse impact on the current family of non-custodial parents or on their participation in the labor market?[128] In order to evaluate the relative importance of these factors, the legislature would need additional empirical research that was designed to evaluate each directly.

Even though the Wisconsin study could not answer the question of whether to adopt wage withholding, it did provide relevant and valuable information. In addition to providing information about the increase in collections due to immediate wage withholding, it also provided information about problems in implementing reforms in the child support system and identified the need for other reforms in the enforcement system if the full potential for child support payments was to be achieved.

---

[127] Id.

[128] For a discussion of a variety of possible negative consequences of aggressive child support enforcement on low-income populations, see Robert F. Kelly, The Family and the Urban Underclass: An Integrative Framework, 6 J. of Fam. Issues 159 (1985).

BEST COPY AVAILABLE

Research can also provide a basis for identifying and questioning normative assumptions that underlie certain law reform proposals. One of the arguments for a primary caretaker presumption, for example, is that child custody claims will be used to threaten primary caretaker parents, typically mothers, into accepting less child support and spousal support than they should receive.

In their indepth, longitudinal study of custody arrangements attending divorce of 1100 California families from two counties, however, Professors Maccoby and Mnookin examined the assumption that fathers use custody to bargain with mothers for reduced financial support in cases in which there was substantial legal conflict. They found no statistically persuasive evidence that this in fact happened. They felt that this result was influenced by the fact that the primary caretaker tended to be awarded custody even though there was no presumption that favored that result. In addition, they thought that the use of child support guidelines and community property rules also limited opportunity to engage in strategic bargaining related to custody and financial issues.[129] Their study also found that parents frequently did not disagree over custody, and for those that did disagree, mediation was beneficial.

Hence the overall picture of what happened in custody

---

[129] For similar results concerning alimony and custody "trade-offs" see Robert F. Kelly & Greer Litton Fox, Determinants of Alimony Awards: An Empirical Test of Current Theories and a Reflection on Public Policy, Syracuse L. Rev. (forthcoming, 1993).

disputes is clarified by the Maccoby and Mnookin study. Their results suggest that rules other than a primary caretaker presumption, such as well defined and enforced child support guidelines and mandatory mediation, might be better ways to achieve the desirable result of keeping custody disputes at a minimum and preventing custody blackmail. The Maccoby and Mnookin study is a valuable challenge to many normative assumptions about the interaction between law and parents' bargaining at divorce, even though the study findings should not be generalized to other states without carefully worded qualification.

Individuals interested in reforming a particular family law area should certainly review policy research that was directed at formulating recommendations in that area. In addition, disciplinary or theory-driven research, as opposed to policy oriented applied research, can provide useful information.[130] In a review of disciplinary research, however, the reviewer not only has to make a judgment about the quality of the research, but also about how relevant the research is to the policy question. The next section examines the function and structure of the research review primarily in the content of policymaking.

B. The Research Review

Research reviews can be used in many different situations by

---

[130] Social science professional associations and legal scholars have used research reviews in amicus briefs. See Roesch, et al., Social Science and the Court: The Role of Amicus Curiae Briefs, 15 L. and Hum. Behav. 1 (1991).

BEST COPY AVAILABLE

72

groups with various goals. They are used by policymakers in medicine in developing standards for clinical practice, health care delivery systems and medical research agendas.[131] They are also used to further disciplinary research. For example, in "pure" disciplinary research it is expected that researchers will assess past research in that area in preparation for undertaking new work. Indeed, an important criterion used to evaluate proposals for research funding by federal agencies such as the National Institutes of Health and the National Science Foundation is the quality of the research review. Publishing research reviews that describe the "state of the field," identify problem areas and encourage new research directions is an important function of scientific journals. Legal scholars may study existing research reviews or perform their own review in order to use social science in their analyses of legal problems. This section will focus on the use of the research review in family policy development, but the principles discussed are applicable to research reviews in general.

The goal of the reviewer is to identify and fairly summarize past research on a particular topic. One way of conceptualizing the reviewer's task is to specify the variables in the studies that are of interest for policy reasons and to consider them in a

---

[131] David B. Larson et al., The Systematic Review: An Innovative Approach to Reviewing Research, 1992, Report prepared for the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services, Washington, DC.

functional model with the form: **Y = f(I,X) + Error.**[132] If the reviewer is interested in the effects of wage withholding on the amount of child support paid, **Y** would represent the outcome or dependent variable of interest, that is, the amount of child support paid. **I** would represent the intervention of interest, that is the policy variable, in the case at hand a new wage withholding system. **X** would represent a set of characteristics of the participants and/or other circumstances that may be related to how the intervention will influence the outcome variable and that, as a result, need to be controlled if we are to understand **I's** true effect on **Y**. Examples of **X's** might be the income of the payor, prior sanctions for nonpayment or the strength of the local labor market. The outcome **Y** depends on the intervention, **I**, and the characteristics of the participants or other circumstances, **X**, plus **Error**, that may be random due to measurement error or sampling error, if a sample is used rather than the entire population, or systematic if the study has significant design flaws.[133]

Using this model, consider the typical questions that the reviewer might ask.[134] First, looking at all of the relevant studies, how does the intervention, **I**, influence the outcome, **Y**, "on average," that is, Does wage withholding tend to increase the amount of child support payments? Second, does the average effect

---

[132] Light & Pillemer, supra note 120 at 113.

[133] Id. at 15-16.

[134] See id.

of the intervention, I, change when conditions specified by the set of variables we have called X are controlled, for example: Is wage withholding more or less effective with high income parents? Does the impact of wage withholding vary depending on the levels of unemployment in the local child support jurisdiction? To answer this question, the reviewer would need to assess how I operates "on average," as well as how I operates jointly with relevant X's such as parental income and prevailing labor market conditions. Third, how well is the intervention likely to work in the particular policy situations of interest? For example, what does the research tell us about how well wage withholding would work in our state? The reviewer would summarize and synthesize the research in an attempt to answer some or all of these questions.

One challenge that the reviewer must confront is how to take into account in the review the fact that studies will vary significantly in their methodological quality, that is, their level of systematic Error. Should weak studies be considered in the review or should they be excluded? If they are considered, what weight should they be given? If a number of low quality studies all point in the same direction and no other research on the topic is available, should these studies be paid substantial attention?[135] The answers to these questions are complex and depend on how the reviewer conceptualizes the relationship

---

[135] See John Eeckelaar & Mavis Maclean, Divorce Law and Empirical Studies- A Reply, 106 L. Q. Rev. 621, 624-25 (1990).

75

between the quality of the research and the findings.

First, how should research quality be assessed? Existing principles of research methodology provide an appropriate basis and framework for judging the quality of social science research. Basic problems related to study design and methodology were discussed in Section II. In addition to these concerns, the interpretation of the data and its analysis, the presentation of findings, the conclusions drawn and the relation of the study to other research should be scrutinized.

Obviously, making these assessments is not an easy task. A researcher typically makes many compromises in the long and complex process of study design, implementation and analysis. Other researchers may feel that some of those choices were wrong and will criticize the research accordingly. Interestingly, as research in an area becomes more sophisticated, evaluating it will also become more complex and will require a greater knowledge of research methodology.[136]

The complexity of the quality assessment in research reviews is illustrated by the problems of evaluation encountered by researchers using a relatively new method for research review and

---

[136] Legal scholars should take heart and realize that even methodologists have a sense of humor. See H. Wayne Hogan, Concentric Zonal Path Analysis of Residual Patterns and Friendship Choices in a Middle-Tennessee County by Sex-Role and Breed of Subject 24 J. of Irreproducible Results 8 (1978)(a fictitious study of the local canine population.)

BEST COPY AVAILABLE

synthesis called meta-analysis.[137] Meta-analysis and related techniques are used to summarize studies quantitatively, rather than through the traditional narrative research review.[138] Meta-analyses begin with the reviewer establishing a comprehensive listing of studies on a particular topic to be reviewed. On-line computer searches greatly facilitate the development of such listings, but recent research suggests that computer searches should not be relied upon exclusively to produce comprehensive listings of research available on a topic.[139] Central to the process of meta-analysis is the development of clear criteria for the inclusion or exclusion of studies in the listing of research to be analyzed. Meta-analysis and related techniques would typically exclude studies that had not appeared in peer review journals or did not report quantitative findings. The rationale for these decision rules are, respectively, that peer review

---

[137] See Light & Pillemer, supra note 120, for the initial conceptualization of meta-analysis. For more recent discussions see Gene V. Glass et al., Meta-Analysis in Social Research (1981) [hereinafter Glass et al]; Robert Rosenthal, Meta-Analytic Procedures for Social Research (1984); and Fredric Wolf, Meta-Analysis: Quantitative Methods for Research Synthesis (1986). Wolf has a summary of criticisms of meta-analysis at 14 and guidelines for practice at 55-56.

[138] A recently developed adaptation of meta-analysis is "systematic review" which has the advantages of being substantially less expensive to conduct than a full-blown meta-analysis and requiring fewer assumptions as to the level of development of the research literature to be reviewed. Because systematic review is closely allied to meta-analysis and in order to limit the length of our discussion, we focus our discussion to meta-analysis, Larson et al., supra note 131.

[139] Joseph C. Bareta et al., A Comparison of Manual and MEDLARS Reviews of the Literature on Consultation-Liason Psychiatry, 44 Am. J. of Psychiatry 1040-42 (1990).

77

BEST COPY AVAILABLE

assures a minimum level of quality and that meta-analysis, as a quantitative technique, requires quantitative data to analyze.[140] But it important to note that criteria such as these have important implications. Qualitative research of the type produced by anthropologists, monographs whose results had not previously appeared in peer-reviewed journals, and law review articles, even if they reported quantitative findings, would be excluded under these rules. Hence, while meta-analysis is a valuable research review technique, important research may be excluded from review.

After listing the studies to be reviewed and the criteria for choosing them, the next step in meta-analysis is to record and statistically analyze the substantive findings and the basic characteristics of the methodological design of each study. If the meta-analyst were interested in the relationship between children's well being and divorce, the findings from all listed studies on this topic would be recorded, but so also would information on each study's design such as the type of sample used, the number of subjects, response rates, reliability of measures and type of controls employed. Both types of information would become variables in analyses to assess if multiple studies demonstrate consistent findings independent of the effects of variation in the quality of research methodology. Obviously, in this endeavor, the meta-analyst would want to systematically apply the same methodological quality criteria to each study

---

[140] See Larson et al., supra note 131, at 8, 10.

included in the meta-analytic review.

Meta-analyst Robert Rosenthal has explained one way of accomplishing this goal. He recommends that all studies used be rated by skilled methodologists.[141] To reduce bias in the ratings, the methodologists should have no special interest in the subject being investigated and should rate the methodology of each study before considering the findings.[142] The methodologists could be asked to judge the research on a number of different levels:

> The specific judgments to ask of our methodologists can range from the most general question of overall quality rated on a 9-point rating scale, to intermediate level questions of quality of design, quality of statistical analysis, quality of ecological validity, and the like, all rated on a 9-point scale, to a series of very specific questions such as : Was random assignment of subjects employed? Was the assumption of independence of errors in the analysis of variance met?[143]

He also recommends that the reliability of the assessors be measured and suggests that the assessors' opinions are unlikely to be unanimous.[144]

Assessing quality with meta-analytic techniques, then, is difficult and costly. Reviews of family research typically have not used this type of rigorously controlled analysis in the assessment of quality, but rather have tended to rely on the

---

[141] Rosenthal, supra note 137, at 55.

[142] Id.

[143] Id.

[144] Id.

individual reviewer's skill alone.[145] Although following Rosenthal's recommendations may not be immediately financially feasible for many reviewers since funding for the outside methodologists suggested by Rosenthal for meta-analysis would be needed, reviewers will increasingly be judged in terms of the steps they have taken to reduce potential sources of bias in their reviews. Indeed, given the increasing acceptance and use of meta-analysis, we suspect that such reviews, in spite of their costs, will become progressively more common in family social science in the coming years and that these reviews will prove extremely valuable for family policy analysis and formation.

There is an important issue raised by our discussion of meta-analysis that should be addressed before leaving the topic of research reviews. Once the quality of studies has been assessed, should a reviewer omit the low quality research from the review? Traditionally, it was assumed that methodological weaknesses would affect the findings and that such studies should be devalued, if not totally discounted in research reviews.

A newer view, supported by meta-analysts, is that whether study quality has influenced the findings should be treated as an

---

[145] See, e.g., narrative reviews such as Glenn, Quantitative Research on Marital Quality in the 1980s: A Critical Review, 52 J. of Marriage and the Fam. 818 (1990), and Marilyn Coleman & Lawrence H. Ganong, Remarriage and Stepfamily Research in the 1980s: Increased Interest in an Old Family Form, 52 J. of Marriage and the Fam. 925 (1990), and other reviews in that same volume reviewing family research in the 1980s. For a quantitative review that relies on the reviewers assessments of quality, see Paul R. Amato and Keith, Parental Divorce and Adult Well-Being: A Meta-Analysis, 53 J. of Marriage and the Fam. 43 (1991).

BEST COPY AVAILABLE

empirical question rather than an a priori assumption.[146] Researchers adopting this approach have analyzed the effect of methodological quality on findings in a number of different research areas that have had both high quality and low quality studies. Surprisingly, their research has shown that quality does not always affect findings, and that in some areas both poor quality and high quality research converge on similar results.[147] Consequently, many meta-analysts feel that before studies are excluded from reviews because of poor methodologies alone, the reviewer should consider and attempt to estimate the actual relationship between design quality and substantive findings. Studies of low quality should be included in meta-analyses and their weaknesses measured and treated as variables. Advocates of meta-analysis feel that the low quality studies can be useful and that "many weak studies can add up to a strong conclusion."[148]

---

[146] Glass et al. supra note 137, at 22.

[147] For example, a review of the research on the effect of deinstitutionalization in mental health concluded that there was no relationship between research design and outcomes. However, a review of the effect of coaching on SAT scores concluded that study findings were affected by research design with observational studies finding that coaching helped a great deal and randomized designs finding that coaching helped very little. Light & Pillemer, supra note 113, at 155-56 (citing studies by Straw, Reinstitutionalization in Mental Health: A Meta-analysis, 8 Evaluation Stud. Rev. -- (1983) and DerSimonian & Laird, Evaluating the Effect of Coaching on SAT Scores: A Meta-analysis, 53 Harv. Educ. Rev. 1 (1983)).

[148] Glass et al. offer the following hypothetical in support of this assertion:

Suppose that, in a group of 100 studies, studies 1-10 are

Note that even if this premise is accepted, a meta-analysis would not necessarily solve difficulties of generalizability in fields where much research shares a common methodological flaw, such as certain areas of family research. Many custody studies, for example, have used samples of middle-income, white families. Even if all of these studies converged on similar findings in a meta-analysis, we could not generalize the results of the analysis to minority, upper-class families if we thought that ethnicity and income were important control variables. Further, it is an assumption of meta-analysis that to determine the relationship between quality and findings, some studies without the particular methodological flaw must be available for purposes of comparison. In an area in which research was relatively new, most or all the studies might share the same basic flaws, such as

---

> weak in representative sampling but strong in other respects; studies 21-30 are weak in internal validity only; studies 31-40 are weak only in data analysis; and so on. But imagine also that all 100 studies are somewhat similar in that they show a superiority of the experimental over the control group. The critic who maintains that the total collection of studies does not support strongly the conclusion of treatment efficacy is forced to invoke an explanation of multiple causality (i.e., the observed difference can be caused either by this particular measurement flaw *or* this particular analysis flaw, *or* . . .). The number of multiple causes which must be invoked to counter the explanation of treatment efficacy can be embarrassingly large for even a few dozen studies. Indeed, the multiple-defects explanation will soon grow into a conspiracy theory or else collapse under its own weight. Respect for parsimony and good sense demands an acceptance of the notion that imperfect studies can converge on a true conclusion.

Glass et al. supra note 137, at 221-22.

BEST COPY AVAILABLE

the use of convenience samples rather than probability samples, and this common flaw will severely limit the reviewer's ability to draw reliable conclusions even with the use of meta-analysis.

Hence, a reviewer, whether planning to use meta-analysis or simply a narrative review, may complete the assessment of the quality of the research in an area of family social science relevant to a legal question and decide that all of the research has substantial and similar design problems because it is primarily exploratory with research questions and methodologies in early stages of development. Is reviewing this incipient research with substantial methodological problems at all useful? We would say yes, but would hasten to emphasize that we do not take this position because we believe that it is possible or wise to rely on such a summary of the research findings in making policy decisions of consequence. Rather, such a review is valuable because it provides information on the state of the disciplinary research and the types of conceptualizations, measurement and hypotheses that are being used. Where findings and conclusions were unexpected, they may help expand our thinking about a particular problem or cause us to question our assumptions.

A major value of a review of incipient research is that it encourages the process of relating and reformulating legal and social science concepts and research. This process, to the degree that it is a cooperative endeavor between the law and social science, is useful both because it can help to clarify legal and

normative assumptions and, reciprocally, because it can provide guidance to the social sciences about the questions that the law needs answered, thereby furthering both disciplinary and policy research. Professor David Chamber's review of disciplinary research related to custody is a good example of this.[149] He reviewed all the existing research that was relevant to the policy questions he was addressing, namely custody disputes involving young children, but found that much of the work was rudimentary and flawed and in some areas there was very little done at all. The review is useful nonetheless to identify where research is needed and how it could inform policy. Chambers was careful to note that the empirical studies, standing alone, did not support a change in existing law.[150]

Although the research review serves a number of useful functions, it should be clear that both narrative and quantitative reviews require a number of personal judgments on the part of the reviewer. Unfortunately, many of the problems that were discussed in Part II about bias in doing research are also present in reviews of the research of others. One of the feminist critiques discussed in that section, that of the anecdoctalists, was addressed to research reviews.[151] The other feminist critiques were not, but their criticism of the conduct

---

[149] David L. Chambers, Rethinking the Substantive Rules for Custody Disputes at Divorce, 83 Mich. L. Rev. 477 (1985).

[150] Id. at 559.

[151] See text accompanying notes 72 & 100-106 infra.

of research would be appropriately directed at research reviews as well. We would support a feminist empiricist view that ideological perspectives may be used to strive for more objectivity in reviews, rather than the anecdoctalist view that research be accepted uncritically or rejected for research reviews based on ideology. Unfortunately, even inadvertent bias can affect decisions about what research to include in research reviews, what weight to give the findings of these studies and how to interpret them. Reviewers have been criticized not only for bias, but also for not being sufficiently systematic and comprehensive. These problems tend to be exacerbated when the number of studies to be reviewed is large.[152] To alleviate these problems the reviewer should clearly indicate the purpose of the review and the methodology used to select studies and measure their substantive findings and methodological characteristics.[153] If research is to be used in policy formation, it should be judged by its quality, including concerns about bias, the relation of quality to findings, and its relevance to the policy question being addressed.[154]

---

[152] See Glass et. al. supra note 137 at 12-13; Light & Pillemer, supra note 120, at 3.

[153] For a helpful checklist for evaluating reviews, see Light & Pillemer, supra note 120, at 160-73.

[154] See Faigman, supra note 5. Faigman recommends that "[t]he legal relevance of social science findings should depend on their scientific strength, that is on the ability of social scientists to answer validly the questions posed to them." Id. at 1009-10. Although we agree that the research should be judged based on scientific principles, we do not think that much family research reaches the level of objectivity that Faigman expects. Hence, we

In summary, the research review is a valuable tool for legal scholars interested in family law analysis and formation. The potential reviewer of family research needs to be aware that it is a difficult task and the technology of the review is developing rapidly. The on-line literature search facilities available today and the advances made in meta-analytic strategies provide major opportunities for collaboration between social scientists and family law analysts. Communication and collaboration across disciplines not only can help family law analysts produce competent reviews of research relevant to their legal problems, but also can help social scientists who typically do not know what legal questions need asking or how to ask them.[155] Family law analysts need to become part of the research review loop. On the other hand, the user of existing research reviews needs to maintain an acute awareness that the research review process, including meta-analysis, has significant limitations and can be poorly handled. Hence, in spite of advances made in the technology of the research review, it is crucial that the legal consumer assume a critical stance whenever using the results of research reviews.

IV. Conclusions and a Proposal

---

would consider the assessment of relevance to be more of a continuum than a dichotomy between "objective" and "not objective."

[155] See e.g. Sarah H. Ramsey, Stepparent Support of Stepchildren: The Changing Legal Context and the Need for Empirical Policy Research, 35 Family Relations 363 (1986).

BEST COPY AVAILABLE

In summary, the research review is a valuable tool for legal scholars interested in family law analysis and formation. The potential reviewer of family research needs to be aware that it is a difficult task and the technology of the review is developing rapidly. The on-line literature search facilities available today and the advances made in meta-analytic strategies provide major opportunities for collaboration between social scientists and family law analysts. Communication and collaboration across disciplines not only can help family law analysts produce competent reviews of research relevant to their legal problems, but also can help social scientists who typically do not know what legal questions need asking or how to ask them.[155] Family law analysts need to become part of the research review loop. On the other hand, the user of existing research reviews needs to maintain an acute awareness that the research review process, including meta-analysis, has significant limitations and can be poorly handled. Hence, in spite of advances made in the technology of the research review, it is crucial that the legal consumer assume a critical stance whenever using the results of research reviews.

IV. Conclusions and a Proposal

---

would consider the assessment of relevance to be more of a continuum than a dichotomy between "objective" and "not objective."

[155] See e.g. Sarah H. Ramsey, Stepparent Support of Stepchildren: The Changing Legal Context and the Need for Empirical Policy Research, 35 Family Relations 363 (1986).

The use of social science research in family law formation is controversial, and the debate over whether or how social science can be used will undoubtedly continue for some time. Although we favor using social science, a major goal of this article was to clarify the issues in the debate. The problems related to the quality of the research and the function of research in policymaking are distinguished in an effort to develop some principled bases for decisions about the use of social science in policymaking.

The article also emphasizes the need for caution in the use of research. One concern is quality. Family research, in part because it is a relatively new and developing field, and in part because its subject matter is a complex and changing social system, has had to deal with many substantive and methodological problems. Although some of these problems, such as the use of small, unrepresentative samples, are becoming less common, others, such as bias, are unavoidable. Another concern is the difficulty in identifying and summarizing relevant research in an accurate and useful way. A final caution is that research cannot replace the normative aspects of decisionmaking. Research can, however, help decisionmakers to be better informed about policy problems and possible solutions. At a minimum, social science research can help us identify our assumptions, and therefore require our policy rationales to be explicitly ideological when they do not have broad empirical support.

As more legal scholars use social science and more social

scientists become familiar with legal issues it will be easier for the disciplines to interact. Family researchers need to be more concerned about the impact of law on families in designing their research. Legal scholars need to be educated and critical reviewers of family research. Dialogue between these disciplines on the relationship of law, social science and families would strengthen both.

To encourage collaboration between lawyers and social scientists with interests in family law formation, we close with a modest but concrete suggestion. We propose that the Family Law Sections of the Association of American Law Schools and the American Bar Association, in cooperation with the family divisions of social science organizations such as the American Sociological Association, the American Psychological Association and the National Council on Family Relations, constitute a panel of lawyers and social scientists concerned with the use of social science research in family law formation. The panel would have two major responsibilities: (1) to develop and issue on a five year cycle calls for social scientific research[156] and, (2) to call for research reviews of existing disciplinary and policy social scientific studies on issues of immediate family policy concern. The premise underlying the formation of such a panel is

---

[156] For an illustration of the type of research that might be called for see Robert F. Kelly, Social Science, Family Law and Divorce: Conceptual and Methodological Reflections on the Need for a New Integration (1993) (paper presented at a seminar sponsored by the Urban Institute's Program of Policy Research on Children and Youth, Washington, D.C.).

that, while it is not necessary for lawyers to become social scientists, or for social scientists to become lawyers, it is crucial for lawyers and social scientists to educate each other and to plan together on a continuing basis. We believe that the analyses we have developed in this paper strongly suggest that there is a pressing need for sustained and systematic efforts to reduce misunderstandings and enhance cooperation between lawyers and social scientists concerned with family law and public policy.

The panel we propose would not be a funding agency, but rather a highly structured forum for lawyers and social scientists to identify their interests to each other and articulate areas of cooperation likely to make significant contributions to the family policy formation process. The calls for studies and research reviews issued by the panel would likely represent an influential voice to government agencies and foundations in their development of research funding priorities. As importantly, such a panel would enhance the likelihood of lawyers getting their questions asked in social science research and social scientists knowing which family law questions need to be asked and how they should be asked. Finally, we suspect that the clearinghouse functions of such a panel would greatly facilitate the formation of lawyer/social scientists teams working on family law issues. Clearly, for an endeavor of this type to achieve its missions, it would be essential that the panel's composition represent a broad spectrum of interests,

BEST COPY AVAILABLE

experiences and perspectives on the law, family systems, and social science.